

# A Flexible Galerkin Scheme for Option Pricing in Lévy Models

Maximilian Gaß and Kathrin Glau

March 29, 2016

Technische Universität München, Center for Mathematics  
kathrin.glau@tum.de, maximilian.gass@mytum.de

## Abstract

One popular approach to option pricing in Lévy models is through solving the related partial integro differential equation (PIDE). For the numerical solution of such equations powerful Galerkin methods have been put forward e.g. by Hilber et al. (2013). As in practice large classes of models are maintained simultaneously, flexibility in the driving Lévy model is crucial for the implementation of these powerful tools. In this article we provide such a flexible finite element Galerkin method. To this end we exploit the Fourier representation of the infinitesimal generator, i.e. the related symbol, which is explicitly available for the most relevant Lévy models. Empirical studies for the Merton, NIG and CGMY model confirm the numerical feasibility of the method.

**Keywords** Lévy processes, partial integro differential equations, pseudo-differential operators, symbol, option pricing, Galerkin approach, finite element method

**Mathematics Subject Classification (2000)** 91G80, 60G51, 35S10, 65M60

## 1 Introduction

In computational finance, methods to solve partial differential equations come into play, when both run time and accuracy matters. In contrast to Monte Carlo for example, run-time is very appealing and a deterministic and conservative error analysis is established and well understood. And, compared to Fourier methods, the possibility to capture path-dependent features like early exercise and barriers is naturally built in. Within these appealing features lies the capacity to attract interest from academia and satisfy the needs of the financial industry alike.

In academia, a series of publications by Cont and Voltchkova (2005b), Hilber, Reich, Schwab and Winter (2009), Salmi, Toivanen and Sydow (2014), Itkin (2015) and Glau (2016b) and the monograph of Hilber, Reichmann, Schwab and Winter (2013) have opened the theory to include even more sophisticated models of Lévy type, resulting in Partial *Integro* Differential Equations (PIDEs). The theoretical results have been validated by sophisticated numerical studies. In this context, Schwab and his working group in particular have taken the lead and unveiled the

potential of PIDE theory for practical purposes in the financial industry. Combining state of the art compression techniques with a wavelet basis finite element setup has resulted in a numerical framework for option pricing in advanced and multivariate jump models and thereby moved academic boundaries.

In the financial industry an awareness of the full potential of these tools is yet to be developed. Advocating the advancement of numerical methods one must acknowledge what practice cherishes most in them. Due to model uncertainty and behavioural characteristics of different portfolios, financial institutions need to deal with a number of different pricing models in parallel. Or, in the words of Föllmer (2009): "In any case, the signal towards the practitioners of risk management is clear: do not commit yourself to a single model, remain flexible, vary the models in accordance with the problem at hand, always keeping in mind the worst case scenario."<sup>1</sup> These features need to be reflected in the numerical environment.

In this article, we aim at reconciling capacities of state of the art P(I)DE tools with the flexibility regarding different model choices as required by industry. Desirable features that such an implementation must offer include

- (1) a degree of accuracy that reaches levels relevant to practical applications and that can be measured and controlled by a theoretical error analysis,
- (2) fast run times,
- (3) low and feasible implementational and maintenance cost,
- (4) a flexibility of the toolbox towards different options and models.

Two standard methods are available for solving P(I)DEs, that is the finite difference approach and the finite element method. More recently, also radial basis methods have been pushed forward to solve pricing PIDEs. In principle these concepts can be implemented in such a way that they achieve the desired features 1.–3. and implementations for a variety of models and option types have already been developed: Finite difference schemes solving PIDEs for pricing European and barrier options with an implementation for Merton and Variance Gamma are provided in Cont and Voltchkova (2005b) and Cont and Voltchkova (2005a). The method has been further developed in different directions, we mention one example, Itkin and Carr (2012), who use a special representation of the equation to derive a finite different scheme for jump diffusions with jump intensity of tempered stable type. Wavelet-Galerkin methods for PIDEs related to a class generalizing tempered stable Lévy processes are derived in Matache et al. (2005) for American options and see e.g. Marazzina et al. (2012) for a high-dimensional extension. Radial basis for the Merton and Kou model, American and European options are provided by Chan and Hubbert (2014) and further developed for CGMY models by Brummelhuis and Chan (2014).

An implementation that is flexible in the driving model as well as in the option type first of all requires a problem formulation covering the collectivity of envisaged models and options. In view of feature (1), a unified approach to the error analysis of the resulting schemes is of equal importance. Galerkin methods, accruing from the Hilbert space formulation of the Kolmogorov equation, seem to be predestined to deliver the adequate level of abstraction for this task. It is this abstract level

---

<sup>1</sup>Translated from German.

that makes Galerkin methods flexible in the option types and the dimensions of the underlying driving process. Consequently, even though Galerkin methods seems more involved at first glance in comparison to finite difference schemes, they still promise to lead to a more lucid code that is easier to maintain and to extend. Another fundamental advantage of Galerkin methods is its theoretical framework that allows for clear and extensive convergence analysis and error estimates, which is of great importance for controlling methodological risk in finance. The finite element, or more general Galerkin methods are therefore our methods of choice.

Unfortunately, even the finite element methods faces numerical challenges when implementing Lévy model based pricing tools. More precisely, the Lévy operator that determines the stiffness matrix is of integro differential type. First, the resulting matrix is densely populated and in general not symmetric. Second, and even more severe, the matrix entries typically are not explicitly available. Instead, they require the evaluation of double integral terms possibly involving a numerically inaccessible Lévy measure. In these cases, a thorough analysis of the respective integrals may lead to approximation schemes deriving the stiffness matrix entries with the required precision. Pursuing this way, however, most likely results in a model specific scheme, contradicting requirement (4).

In this article, our aim is *the development of a model-independent approach to set up a FEM solver for option pricing in Lévy models* that we call *symbol method*. We address this goal by expressing the operator in the Fourier space. This means accessing the model specific information via the symbol. In contrast to the operator, the symbol is explicitly available for a variety of models and is thus numerically accessible. Further advantages will be highlighted in subsequent sections.

Section 2 introduces the theoretical framework for our PIDEs of interest and their weak formulation. The next section describes the solution scheme, that is the Galerkin approximation in space. We investigate the scheme with regard to the numerical challenges arising during its implementation. Section 4 introduces the symbol method itself. All components of the FEM solver are expressed in Fourier space. The subsequent numerical evaluation of the stiffness matrix entries is supported by an elementary approximation result. Several examples of symbols for well-known Lévy models confirm the wide applicability of the method and its numerical advantages. Two proposals for the implementation of basis functions are presented. The numerical studies in Section 5 confirm theoretical prescribed rates of convergence and validate the claim of numerical feasibility.

## 2 Kolmogorov equations for option pricing in Lévy models

We first introduce the underlying stochastic processes, the Kolmogorov equation, its weak formulation as well as the solution spaces of our choice.

### 2.1 Lévy processes

Let a stochastic basis  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, P)$  be given and let  $L$  be an  $\mathbb{R}^d$ -valued *Lévy process* with characteristics  $(b, \sigma, F; h)$ , i.e. for fixed  $t \geq 0$  its characteristic function

is given by

$$E e^{i\langle \xi, L_t \rangle} = e^{-tA(-i\xi)} \quad \text{for every } \xi \in \mathbb{R}^d, \quad (1)$$

where the *symbol of the process* is defined as

$$A(\xi) := \frac{1}{2} \langle \xi, \sigma \xi \rangle + i \langle \xi, b \rangle - \int_{\mathbb{R}^d} \left( e^{-i\langle \xi, y \rangle} - 1 + i \langle \xi, h(y) \rangle \right) F(dy). \quad (2)$$

Here,  $\sigma$  is a symmetric, positive semi-definite  $d \times d$ -matrix,  $b \in \mathbb{R}^d$ , and  $F$  is a Lévy measure, i.e. a positive Borel measure on  $\mathbb{R}^d$  with  $F(\{0\}) = 0$  and  $\int_{\mathbb{R}^d} (|x|^2 \wedge 1) F(dx) < \infty$ . Moreover,  $h$  is a truncation function i.e.  $h : \mathbb{R}^d \rightarrow \mathbb{R}$  such that  $\int_{\{|x|>1\}} h(x) F(dx) < \infty$  with  $h(x) = x$  in a neighbourhood of 0. The *Kolmogorov operator of a Lévy process*  $L$  with characteristics  $(b, \sigma, F; h)$  is given by

$$\begin{aligned} \mathcal{A} \varphi(x) := & -\frac{1}{2} \sum_{j,k=1}^d \sigma^{j,k} \frac{\partial^2 \varphi}{\partial x_j \partial x_k}(x) - \sum_{j=1}^d b^j \frac{\partial \varphi}{\partial x_j}(x) \\ & - \int_{\mathbb{R}^d} \left( \varphi(x+y) - \varphi(x) - \sum_{j=1}^d \frac{\partial \varphi}{\partial x_j}(x) h_j(y) \right) F(dy) \end{aligned} \quad (3)$$

for every  $\varphi \in C_0^\infty(\mathbb{R}^d)$ , where  $h_j$  denotes the  $j$ -th component of the truncation function  $h$ .

## 2.2 Kolmogorov equation in variational form

Key for the variational formulation of Kolmogorov equation

$$\partial_t u + \mathcal{A} u = f \quad (4)$$

$$u(0) = g \quad (5)$$

is the definition of the bilinear form

$$a(\varphi, \psi) := \int_{\mathbb{R}^d} (\mathcal{A} \varphi)(x) \psi(x) dx \quad \text{for all } \varphi, \psi \in C_0^\infty(\mathbb{R}^d). \quad (6)$$

It is one of the major advantages of variational formulations of evolution equations that solution spaces of low regularity, as compared to the space  $C^2$  for example, are incorporated in an elegant way. Departing from the space  $C_0^\infty(\mathbb{R}^d)$  of smooth functions with compact support, we can select from a large variety of function spaces  $V$  that is characterized by the following assumption.

(A1)  $V$  and  $H$  are Hilbert spaces such that  $C_0^\infty(\mathbb{R}^d)$  lies dense in  $V$  and there exists a continuous embedding from  $V$  into  $H$ .

Existence and uniqueness of a variational solution critically hinges on the following two properties of the bilinear form:

(A2) *Continuity*: There exists a constant  $C > 0$  such that

$$|a(\varphi, \psi)| \leq C \|\varphi\|_V \|\psi\|_V \quad \text{for all } \varphi, \psi \in C_0^\infty(\mathbb{R}^d).$$

(A3) *Gårding inequality*: There exists constants  $G > 0$  and  $G' \geq 0$  such that

$$a(\varphi, \varphi) \geq G\|\varphi\|_V^2 - G'\|\varphi\|_H^2 \quad \text{for all } \varphi \in C_0^\infty(\mathbb{R}^d).$$

We observe that due to (A1) and (A2), the bilinear form  $a$  possesses a unique continuous bilinear extension  $a : V \times V$  that is continuous, i.e. for a constant  $C > 0$  we have  $|a(\varphi, \psi)| \leq C\|\varphi\|_V\|\psi\|_V$  for all  $v \in V$ . Also (A3) holds for all  $v \in V$ .

As  $V$  is separable, this is also true for  $H$  and one can find a continuous embedding from  $H$  to the dual space  $V^*$  of  $V$ , i.e.  $(V, H, V^*)$  is a Gelfand triplet. We then denote by  $L^2(0, T; H)$  the space of weakly measurable functions  $u : [0, T] \rightarrow H$  with  $\int_0^T \|u(t)\|_H^2 dt < \infty$  and by  $\partial_t u$  the derivative of  $u$  with respect to time in the distributional sense. For a detailed definition which relies on the Bochner integral we refer to Section 24.2 in Wloka (1987). The Sobolev space

$$W^1(0, T; V, H) := \left\{ u \in L^2(0, T; V) \mid \partial_t u \in L^2(0, T; V^*) \right\}, \quad (7)$$

will play the role of the solution space in the variational formulation of Kolmogorov equation (4), (5).

**Definition 2.1.** *Let  $f \in L^2(0, T; V^*)$  and  $g \in H$ . Then  $u \in W^1(0, T; V, H)$  is a variational solution of Kolmogorov equation (4), if for almost every  $t \in (0, T)$ ,*

$$\langle \partial_t u(t), v \rangle_H + a(u(t), v) = \langle f(t), v \rangle_{V^* \times V} \quad \text{for all } v \in V \quad (8)$$

*and  $u(t)$  converges to  $g$  for  $t \downarrow 0$  in the norm of  $H$ .*

**Remark 2.2.** *Assumptions (A1)–(A3) guarantee the existence and uniqueness of a variational solution  $u \in W^1(0, T; V, H)$  of (8), see for instance Theorem 23.A in Zeidler (1990).*

### 2.3 Solution spaces

Definition 6 is based on the  $L^2$ -scalar product and is appropriate for variational equations in Sobolev spaces. Then, typically  $H = L^2$ . For Kolmogorov equations for option prices the initial condition  $g$  in (5) plays the role of the (logarithmically transformed) payoff function of the option. For a call option with strike  $K$  it is of the form  $x \mapsto (S_0 e^x - K)^+$ , for a digital up and out option it is given by  $x \mapsto \mathbf{1}_{x < b}$  with for some  $b \in \mathbb{R}$ . We thus have to observe that the initial condition  $g$  is not square integrable for most of the typical cases of interest. Therefore, we base our analysis more generally on exponentially weighted  $L^2$  spaces: For  $\eta \in \mathbb{R}^d$  let

$$L_\eta^2(\mathbb{R}^d) := \left\{ u \in L_{loc}^1(\mathbb{R}^d) \mid u e^{\langle \eta, \cdot \rangle} \in L^2(\mathbb{R}^d) \right\}, \quad \|u\|_{L_\eta^2} := \left( \int_{\mathbb{R}^d} |u(x)|^2 e^{2\langle \eta, x \rangle} dx \right)^{1/2}$$

and

$$a(\varphi, \psi) := \langle \mathcal{A}\varphi, \psi \rangle_{L_\eta^2} = \int_{\mathbb{R}^d} (\mathcal{A}\varphi)(x) \psi(x) e^{2\langle \eta, x \rangle} dx \quad \text{for all } \varphi, \psi \in C_0^\infty(\mathbb{R}^d). \quad (9)$$

We notice that all assertions of the precedent section, concerning assumptions (A)–(A3) and variational equations hold for bilinear form  $a$  defined by (9) instead of  $a$  from (6) as well.

As solution spaces  $V$  we consider weighted Sobolev-Slobodeckii spaces. These have proven to apply to a large set of option types and models. We refer to Eberlein and Glau (2014) and Glau (2016b), where particularly Feynman-Kac type formulas have been derived linking European and path-dependent options to weak solutions of Kolmogorov equations in Sobolev-Slobodeckii spaces.

To introduce the spaces, we denote by  $C_0^\infty(\mathbb{R}^d)$  the set of smooth real-valued functions with compact support in  $\mathbb{R}^d$  and let

$$\mathcal{F}(\varphi)(\xi) := \int_{\mathbb{R}^d} e^{i\langle \xi, x \rangle} \varphi(x) dx \quad (10)$$

be the Fourier transform of  $\varphi \in C_0^\infty(\mathbb{R}^d)$  and  $\mathcal{F}^{-1}$  be its inverse. We define the *exponentially weighted Sobolev-Slobodeckii space*  $H_\eta^\alpha(\mathbb{R}^d)$  with index  $\alpha \geq 0$  and weight  $\eta \in \mathbb{R}^d$  as the completion of  $C_0^\infty(\mathbb{R}^d)$  with respect to the norm  $\|\cdot\|_{H_\eta^\alpha}$  given by

$$\|\varphi\|_{H_\eta^\alpha}^2 := \int_{\mathbb{R}^d} (1 + |\xi|)^{2\alpha} |\mathcal{F}(\varphi)(\xi - i\eta)|^2 d\xi. \quad (11)$$

Furthermore, we denote the dual space of  $H_\eta^\alpha(\mathbb{R}^d)$  by  $(H_\eta^\alpha(\mathbb{R}^d))^*$ .

### 3 Implementational Challenges

#### 3.1 Abstract Galerkin approximation in space

For a countable Riesz basis  $\{\varphi_1, \varphi_2, \dots\}$  of  $V$  we define

$$V_N := \text{span}\{\varphi_1, \dots, \varphi_N\} \quad \text{for all } N \in \mathbb{N}.$$

Since  $V$  is dense in  $H$ , we may further choose  $g_N$  in  $V_N$  such that  $g_N \rightarrow u(0)$  in  $H$ . For each fixed  $N \in \mathbb{N}$  the semidiscrete problem is defined by restricting (8) to the finite dimensional space: *Find a function  $v_N \in W^1(0, T; V_N; H \cap V_N)$  that satisfies for all  $\chi \in C_0^\infty(0, T)$  and  $\varphi \in V_N$ ,*

$$-\int_0^T \langle v_N(t), \varphi \rangle_{L^2} \dot{\chi}(t) dt + \int_0^T a(v_N(t), \varphi) \chi(t) dt = \int_0^T \langle f(t)|\varphi \rangle_{V^* \times V} \chi(t) dt \quad (12)$$

$$v_N(0) = g_N.$$

As a result of the elegant Hilbert space formulation, the semidiscrete problem (12) is uniquely solvable and the convergence of the sequence  $v_N$  to  $v$  is guaranteed, see Theorem 23.A. and Remark 23.25 in Zeidler (1990).

The major advantage of equation (12) in regard to implementation is that it suffices to insert the basis functions as test functions. Thus, denoting  $g_N = \sum_{k=1}^N \alpha_k \varphi_k$  and  $v_N(t) := \sum_{k=1}^N V_k(t) w_k$  we arrive at

$$\sum_{k=1}^N \dot{V}_k(t) \langle \varphi_k, \varphi_j \rangle_{L^2} + \sum_{k=1}^N V_k(t) a(\varphi_k, \varphi_j) = \langle f(t)|\varphi_j \rangle_{V^* \times V}$$

$$V_k(0) = \alpha_k \quad \text{for all } k = 1, \dots, N.$$

Written in matrix form the problem is to find  $V : [0, T] \rightarrow \mathbb{R}^N$  such that

$$M\dot{V}(t) + AV(t) = F(t) \quad (13)$$

$$V(0) = \alpha, \quad (14)$$

where  $F = (F_1, \dots, F_N)^\top$  with  $F_k(t) = \langle f(t) | \varphi_k \rangle_{V^* \times V}$  for  $k = 1, \dots, N$ ,  $\alpha = (\alpha_1, \dots, \alpha_N)^\top$ , and the *mass matrix*  $M$  and *stiffness matrix*  $A$  are given by

$$M_{jk} = \langle \varphi_k, \varphi_j \rangle_{L^2}, \quad A_{jk} = a(\varphi_k, \varphi_j) \quad \text{for all } j, k = 1, \dots, N. \quad (15)$$

### 3.2 Flexible implementation for different driving Lévy processes

We inspect equation (14) in regard to flexibility towards different options as well as models. All ingredients in (13) depend on the choice of the basis. While  $M$  is independent of the specific problem at hand,  $F$  and  $\alpha$  represent the input data and therefore may vary for different option types. The stiffness matrix  $A$  that carries the information of the driving driving process. So in order to obtain flexibility towards model types, we need a generic way to compute the entries of the stiffness matrix. For smooth basis functions with compact support and solution spaces without weighting, i.e.  $\eta = 0$ , according to (2), (6), the stiffness matrix entries are given by

$$\begin{aligned} a(\varphi_k, \varphi_j) = & - \sum_{l,m=1}^d \frac{\sigma^{j,m}}{2} \int_{\mathbb{R}^d} (\partial_l \partial_m \varphi_k(x)) \varphi_j(x) dx - \sum_{l=1}^d b^l \int_{\mathbb{R}^d} \partial_l \varphi_k(x) \varphi_j(x) dx \\ & - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left( \varphi_k(x+y) - \varphi_k(x) - \sum_{l=1}^d \partial_l \varphi_k(x) h_l(y) \right) F(dy) \varphi_j(x) dx. \end{aligned} \quad (16)$$

Typical basis functions are not smooth. Therefore it is not a priori clear if the integral representation (16) extends to the usual basis functions. Observe that an extension of this representation requires some care: For a large and important class of pure jump Lévy processes, the solution spaces are Sobolev-Slobodeckii spaces of fractional order, i.e.  $H^\alpha$  with some  $0 < \alpha < 1$ . For functions in  $H^\alpha$  with  $\alpha < 1$ , however, the first order weak derivative in (16) is not defined and therewith this integral representation of the bilinear form is not well-defined. Seen that the basis functions usually are in  $H^1$ , also in the more challenging case of solution spaces with fractional order derivatives we derive the validity of the representation under appropriate assumptions.

**Lemma 3.1.** *Let  $d = 1$ . Let  $a$  be defined by (9). Assume (A1)–(A3) for  $a$ ,  $V$  and  $H$  and denote by  $a : V \times V$  its unique bilinear continuous extension. If  $H_\eta^1(\mathbb{R}) \subset V$ , we have for every  $\varphi, \psi \in H_\eta^1(\mathbb{R})$ ,*

$$\begin{aligned} a(\varphi, \psi) = & \frac{\sigma}{2} \int_{\mathbb{R}} \varphi'(x) \psi'(x) e^{2\eta x} dx - b(\eta, \sigma, F) \int_{\mathbb{R}} \varphi'(x) \psi(x) e^{2\eta x} dx \\ & - \int_{\mathbb{R}} \int_{|y| < 1} \int_0^y \int_0^z \varphi'(x+v) dv dz F(dy) (\psi'(x) + 2\eta \psi(x)) e^{2\langle \eta, x \rangle} dx \\ & - \int_{\mathbb{R}} \int_{|y| > 1} (\varphi(x+y) - \varphi(x)) F(dy) \psi(x) e^{2\langle \eta, x \rangle} dx \end{aligned} \quad (17)$$

with

$$b(\eta, \sigma, F) = b - 2\sigma\eta + \int_{|y|<1} (y - h(y))F(dy) - \int_{|y|>1} h(y)F(dy).$$

The proof is provided in Section A.1.

Inspecting the expression for the bilinear form we encounter several numerical challenges due to the integral part—stemming from the jumps of the process:

1. The appealing tridiagonal structure of the stiffness matrix related to the Black-Scholes equation does not extend to the general Lévy setting. Instead, the stiffness matrix is densely populated. Pleasantly, it is still a Toeplitz matrix.
2. For some choices of Lévy measures and bases the stiffness matrix entries may be derived in closed form. This is for instance the case for the Merton model and piecewise linear basis functions. Following Section 10.6.2 in Hilber et al. (2013), the stiffness matrix entries may be derived in semi-closed form expressions for a further group of jump intensities including tempered stable, CGMY and KoBoL processes and the choice of piecewise linear basis functions.

An implementation that is flexible in the driving Lévy process therefore has to rely on numerical approximations of the entries of the stiffness matrix. These approximations inevitably affect the accuracy of the solution to the scheme (13)–(14). The question rises: *How accurate does the integration routine have to be chosen in order to meet a desired accuracy of the solution  $V$ ?* In order to gain a first practical insight in the magnitude of the error resulting from an inaccuracy in the stiffness matrix entries, we next conduct an empirical investigation.

### 3.3 An accuracy study for approximations of the stiffness matrix

To this end we conduct an accuracy study for the stiffness matrix entries simulating the propagation of integration errors in the stiffness matrix up until computed option prices. We choose the Black-Scholes model for which the precise values of all matrix entries are known and assume a market volatility of  $\sigma = 0.2$  and interest rate  $r = 0.01$  for this study. With a classic FEM solver we price a put option with strike  $K = 1$  and maturities  $T \in [0, 3]$  for current values of the stock  $S_0 \in [S_{\min}, S_{\max}]$  with  $S_{\min} = 0.01$  and  $S_{\max} = 10$ . We set the number of involved FEM hat functions to  $N = 150$ , resulting in a mesh with 150 inner grid nodes and mesh fineness  $h = 0.0464$ . We know the mass matrix of the Black-Scholes model to be given by

$$M = \frac{h}{6} \begin{pmatrix} 4 & 1 & 0 & \cdots & 0 \\ 1 & 4 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & 4 & 1 \\ 0 & \cdots & 0 & 1 & 4 \end{pmatrix} \in \mathbb{R}^{N \times N},$$

and the stiffness matrix to be given by

$$A = A^{\text{bs}} = A^{(1)} + A^{(2)} + rM \in \mathbb{R}^{N \times N}, \quad (18)$$



where

$$A^{(1)} = \frac{1}{2} \left( r - \frac{\sigma^2}{2} \right) \begin{pmatrix} 0 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & 0 & -1 \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}, \quad A^{(2)} = \frac{\sigma^2}{2} \frac{1}{h} \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix}.$$

With these matrices we set up a theta scheme,  $\theta = 0.5$ , and derive Black-Scholes put option prices. We use these matrices to simulate how the resulting pricing surface is affected by inaccuracies that might occur when these integrals are solved numerically, instead. To this extent we take the correct stiffness matrix given by (18) and distort each of its entries randomly at different positions  $D \in \mathbb{N}$  after the decimal point by adding  $\varepsilon_i^D = 10^{-(D-1)}\varepsilon_i$  with random  $\varepsilon_i \in (-1, 1)$  for  $i \in \{-(N-1), \dots, -1, 0, 1, \dots, (N-1)\}$  onto the (side) diagonal  $i$  of Matrix  $A$ . Each individual (side) diagonal of the original stiffness matrix is thus affected evenly, keeping the Toeplitz structure of the matrix intact. Since the value of  $A_{ij}$  is only determined by the value of  $j - i$ , this distortion mimics the influence that integration inaccuracies would have.

Hence, for  $D \in \mathbb{N}$  we define the distorted stiffness matrix by

$$A_{\text{distort}}^D = A + \varepsilon^D \in \mathbb{R}^{N \times N}, \quad (19)$$

with

$$\varepsilon^D = 10^{-(D-1)} \begin{pmatrix} \varepsilon_0 & \varepsilon_1 & \varepsilon_2 & \cdots & \cdots & \cdots & \varepsilon_{N-1} \\ \varepsilon_{-1} & \varepsilon_0 & \varepsilon_1 & \ddots & \ddots & \ddots & \vdots \\ \varepsilon_{-2} & \varepsilon_{-1} & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \varepsilon_1 & \varepsilon_2 \\ \vdots & \ddots & \ddots & \ddots & \varepsilon_{-1} & \varepsilon_0 & \varepsilon_1 \\ \varepsilon_{-(N-1)} & \cdots & \cdots & \cdots & \varepsilon_{-2} & \varepsilon_{-1} & \varepsilon_0 \end{pmatrix} \in \mathbb{R}^{N \times N},$$

with uniformly distributed  $\varepsilon_i \in (-1, 1)$ ,  $i \in \{-(N-1), \dots, -1, 0, 1, \dots, (N-1)\}$ , that are drawn independently from each other. Using these distorted stiffness matrices  $A_{\text{distort}}^D$  for different values  $D \in \mathbb{N}$ , we derive again price surfaces of the put option in the Black-Scholes model and compare the difference between the prices coming from the distorted stiffness matrix  $A_{\text{distort}}^D \in \mathbb{R}^{N \times N}$  to the prices from the intact stiffness matrix  $A \in \mathbb{R}^{N \times N}$ .

Figure 1 shows the results of the study and emphasizes the necessity for high numerical accuracy in the computation of the stiffness matrix. We observe that the absolute price differences decrease almost linearly in  $D$ . An accuracy of  $D = 3$  corresponds to integration results that are exact up to the third digit after the decimal point. Pricing resulting from stiffness matrices computed with such a low integration accuracy are unacceptable. The respective pricing errors observable in the top left corner of Figure 1 indicate relative errors of several hundred percent points. With more precise integration results, the error decreases in  $D$  until highly

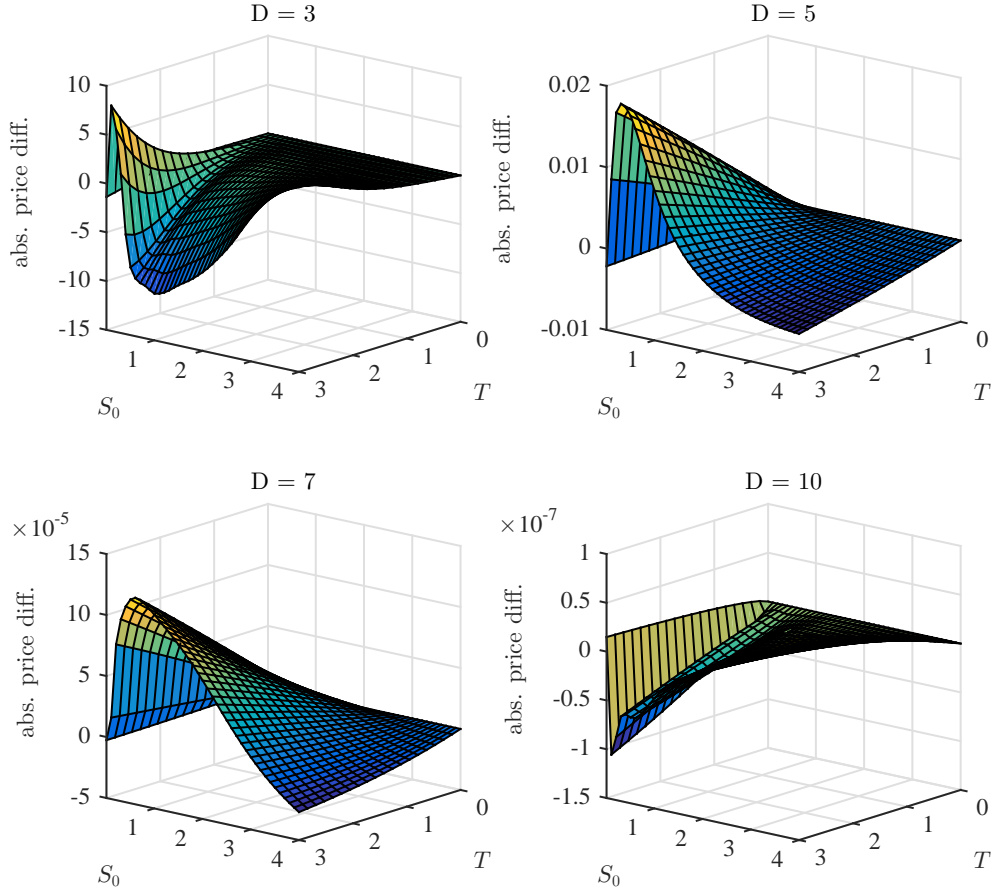


Figure 1: Absolute price differences resulting from a distortion of the stiffness matrix  $A$ . True and distorted prices describe the market value of a put option in the Black-Scholes model with the parametrization from Section 3.3. We compare the price surfaces coming from a theta scheme using the stiffness matrix  $A$  given by (18) to the respective pricing surface when  $A$  is replaced by  $A_{\text{distort}}^D$ , the distorted version of  $A$  as defined in (19), for different values of  $D \in \mathbb{N}$ . The influence of the distortion decreases in  $D$ .

appealing pricing results are achieved for  $D = 7$  and beyond. The magnitude of the pricing error resulting from a distorted stiffness matrix emphasizes the necessity of being able to derive the stiffness matrix entries as accurately as possible.

## 4 Fourier approach to the Kolmogorov equation

In regard to the high accuracy the approximation of the stiffness matrix entries needs to achieve, we would like to avoid numerical evaluations of the stiffness matrix entries on the basis of representation (17). Seeking for alternative representations of the stiffness matrix, let us point out that the symbol  $A$  of the Lévy process is always available. Even more, it is an explicit function of the parameter of the process and thus can be seen as the modelling quantity of the process as the examples 4.6–4.9 below show. We therefore take a Fourier perspective on the variational formulation of the Kolmogorov equation. This is especially promising since the Kolmogorov operator  $\mathcal{A}$  of a Lévy process is a pseudo differential operator with symbol  $A$ ,

$$\mathcal{A}\varphi = \mathcal{F}^{-1}(A\mathcal{F}(\varphi)) \quad \text{for all } \varphi \in C_0^\infty(\mathbb{R}^d), \quad (20)$$

as elementary manipulations show. Now Parseval's identity yields

$$a(\varphi, \psi) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \mathcal{F}(\mathcal{A}\varphi)(\xi) \overline{\mathcal{F}(\psi)(\xi)} d\xi \quad (21)$$

for all  $\varphi, \psi \in C_0^\infty(\mathbb{R}^d)$ , respectively,

$$a(\varphi, \psi) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} A(\xi) \mathcal{F}(\varphi)(\xi) \overline{\mathcal{F}(\psi)(\xi)} d\xi. \quad (22)$$

This well-known identity has already proven highly beneficial for the analysis of the variational solutions of the Komogorov equations, compare e.g. Hilber et al. (2013), Glau (2016a) and Glau (2016b). Here we exploit this representation for the numerical implementation.

**Lemma 4.1** (Continuous extension of bilinear forms). *Let  $A$  be the symbol of a Lévy process given by the characteristic triplet  $(b, \sigma, F)$ . Denote by  $\mathcal{A} : C_0^\infty(\mathbb{R}^d, \mathbb{C}) \rightarrow C^\infty(\mathbb{R}^d, \mathbb{C})$  the pseudodifferential operator associated with symbol  $A$ . Furthermore, denote by  $a : C_0^\infty \times C_0^\infty \rightarrow \mathbb{C}$  the bilinear form associated with the operator  $\mathcal{A}$ . Let  $\eta \in \mathbb{R}^d$ . If*

*i) the exponential moment condition*

$$\int_{|x|>1} e^{-\langle \eta', x \rangle} F(dx) < \infty \quad (23)$$

*holds for all  $\eta' \in \text{sgn}(\eta^1)[0, |\eta^1|] \times \cdots \times \text{sgn}(\eta^d)[0, |\eta^d|]$  and*

*ii) there exists a constant  $C_1 > 0$  with*

$$|A(z)| \leq C_1(1 + \|z\|)^\alpha \quad (24)$$

*for all  $z \in U_{-\eta}$  where*

$$U_{-\eta} = U_{-\eta^1} \times \cdots \times U_{-\eta^d} \quad (25)$$

*with  $U_{-\eta^j} = \mathbb{R} - i \text{sgn}(\eta^j)[0, |\eta^j|]$ ,*

then  $a(\cdot, \cdot)$  possesses a unique linear extension  $a : H_\eta^{\alpha/2} \times H_\eta^{\alpha/2} \rightarrow \mathbb{C}$  which can be written as

$$a(\varphi, \psi) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} A(\xi - i\eta) \widehat{\varphi}(\xi - i\eta) \overline{\widehat{\psi}(\xi - i\eta)} d\xi \quad (26)$$

for all  $\varphi, \psi \in H_\eta^{\alpha/2}(\mathbb{R}^d)$ .

*Proof.* The proof can be found in Eberlein and Glau (2011) using Theorem 4.1 therein and Parseval's identity.  $\square$

In order to gain first insight in the convergence analysis, we fix a level  $N$  in the Galerkin scheme and derive conditions for the convergence of the sequence of weak solutions that we obtain by approximating the stiffness matrix entries. In the implementation in Section 5 below we will also approximately compute the right hand side of the equation. We therefore more generally consider sequences of stiffness matrices, right hand sides and initial conditions.

As usual, we denote for a given bilinear form  $a : V \times V \rightarrow \mathbb{R}$  the associated operator  $\mathcal{A} : V \rightarrow V^*$  defined by  $\mathcal{A}(u)(v) := a(u, v)$  for all  $u, v \in V$ .

**Lemma 4.2.** *Let  $V$ ,  $H$  and  $a : V \times V \rightarrow \mathbb{R}$  satisfy (A1)–(A3). Let  $X := \text{span}\{\varphi_1, \dots, \varphi_N\} \subset V$  and for each  $n \in \mathbb{N}$  let*

(An1)  $f_n, f \in L^2(0, T; H)$  with  $f_n \rightarrow f$  in  $L^2(0, T; X^*)$ ,

(An2)  $g_n, g \in H$  with  $g_n \rightarrow g$  in  $H$ ,

(An3)  $a_n : V \times V \rightarrow \mathbb{R}$  be a bilinear form such that for all  $j, k \leq N$ ,

$$|(a_n - a)(\varphi_j, \varphi_k)| \rightarrow 0. \quad (27)$$

Then the sequence of unique weak solutions  $v_n \in W^1(0, T; X, H)$  of

$$\dot{v}_n + \mathcal{A}_n v_n = f_n, \quad v_n(0) = g_n \quad (28)$$

converges strongly in  $L^2(0, T; X) \cap C(0, T; H)$  to the unique weak solution  $v \in W^1(0, T; X, H)$  of

$$\dot{v} + \mathcal{A} v = f, \quad v(0) = g. \quad (29)$$

The proof is provided in Section A.2.

Profound convergence analysis providing asymptotic convergence rates for the fully discrete scheme can be derived based on the techniques presented by von Petersdorff and Schwab (2003), which is beyond the scope of the present article.

## 4.1 The symbol method

The key component of a Galerkin FEM solver is the model dependent stiffness matrix. Using expression (16) of section 3.2 above, the entries of that matrix can be derived. The existence of the Lévy measure  $F$  in that expression, however, renders the numerical derivation of the matrix rather cumbersome. Additionally, the empirical accuracy study of section 3.3 emphasize that utmost care must be taken when the stiffness matrix entries are numerically derived. Consequently, in this section we approach the calculation of that FEM solver components differently.

The Fourier approach indicated by Lemma 4.1 will allow us to access the model information required for the stiffness matrix and all other FEM solver components not via the operator but rather via the associated symbol, instead. In stark contrast to the operator, the symbol of a Lévy model is numerically accessible in a unified way for a large set of underlying models and we will present several examples.

Let us state the core lemma of this section. Here we concentrate on basis functions obeying a simple nodal translation property, which is in particular satisfied for the classical piecewise polynomial basis functions. The translation property can also easily be extended to that one satisfied by biorthogonal wavelet bases that haven't proved to be useful for solving Komlogorov PIDEs for option pricing, compare Matache et al. (2004).

**Lemma 4.3** (Symbol method for bilinear forms). *Let the assumptions of Lemma 4.1 be satisfied with  $\eta = 0$ . Assume further for  $N \in \mathbb{N}$  a set of functions  $\varphi_0, \varphi_1, \dots, \varphi_N \in H_0^{\alpha/2}(\mathbb{R})$  and nodes  $x_1, \dots, x_N \in \mathbb{R}$ , such that for all  $j = 1, \dots, N$*

$$\varphi_j(x) = \varphi_0(x - x_j), \quad \forall x \in \mathbb{R},$$

*holds. Then we have*

$$a(\varphi_l, \varphi_k) = \frac{1}{2\pi} \int_{\mathbb{R}} A(\xi) e^{i\xi(x_l - x_k)} |\widehat{\varphi_0}(\xi)|^2 d\xi. \quad (30)$$

*for all  $k, l = 1, \dots, N$ . If additionally*

$$\Re(A(\xi)) = \Re(A(-\xi)) \quad \text{and} \quad \Im(A(\xi)) = -\Im(A(-\xi)), \quad (31)$$

*then*

$$a(\varphi_l, \varphi_k) = \frac{1}{\pi} \int_0^\infty \Re \left( A(\xi) e^{i\xi(x_l - x_k)} \right) |\widehat{\varphi_0}(\xi)|^2 d\xi \quad (32)$$

*for all  $k, l = 1, \dots, N$ .*

*Proof.* Elementary properties of the Fourier transform yield

$$\widehat{\varphi_j}(\xi) = e^{i\xi x_j} \widehat{\varphi_0}(\xi), \quad \forall \xi \in \mathbb{R}, \quad (33)$$

where

$$\widehat{\varphi_0}(\xi) = \frac{2}{\xi^2 h} (1 - \cos(\xi h)), \quad \forall \xi \in \mathbb{R}. \quad (34)$$

Since  $\varphi_j \in H_0^{\alpha/2}(\mathbb{R})$ , for all  $j = 1, \dots, N$ , the identity (30) follows from Lemma 4.3 above or Theorem 4.1 and Remark 5.2 and the lines thereafter in Eberlein and Glau (2011), respectively. The second claim (32) is then elementary.  $\square$

**Corollary 4.4** (Symbol method for stiffness matrices). *Let  $A \in S_\alpha^0$  be a univariate symbol with associated operator  $\mathcal{A}$ . Denote by  $\varphi_j \in L^1(\mathbb{R})$ ,  $j = 1, \dots, N$  the basis functions of a Galerkin scheme associated with an equidistantly spaced grid  $\Omega = \{x_1, \dots, x_N\}$  possessing the property*

$$\varphi_j(x) = \varphi_0(x - x_j), \quad \forall x \in \mathbb{R}, \quad (35)$$

*for some  $\varphi_0 : \mathbb{R} \rightarrow \mathbb{R}$ . Then, the stiffness matrix  $A \in \mathbb{R}^{N \times N}$  of the scheme can be computed by*

$$A_{kl} = \frac{1}{2\pi} \int_{\mathbb{R}} A(\xi) e^{i\xi(x_l - x_k)} |\widehat{\varphi_0}(\xi)|^2 d\xi \quad (36)$$

*for all  $k, l = 1, \dots, N$ .*

*Proof.* The proof is an immediate consequence of Lemma 4.3.  $\square$

**Remark 4.5** (On the symbol method for bilinear forms). *From a numerical perspective, the representations of the stiffness matrix entries provided in Lemma 4.3 and Corollary 4.4 are highly promising:*

1. *Instead of the double integrals appearing in (16), only one dimensional integral need to be computed.*
2. *The model specific information is expressed via the symbol  $\xi \mapsto A(\xi)$ , which for a large set of models is available in form of an explicit function of  $\xi$  and the model parameters, a feature that we now can exploit numerically. We give a short list of examples below. For further examples we refer to Glau (2016a) and Glau (2016b).*
3. *Representation (36) displays the entries of the stiffness matrix as Fourier integrals. Moreover, the nodes appear as Fourier variables. As a consequence, Fast Fourier Transform (FFT) methods can be used to accelerate their simultaneous computation.*

Expression (3) introduces operators  $\mathcal{A}$  for Lévy processes  $L$  in terms of the characteristic triplet  $(b, \sigma, F)$ . The following examples present the respective symbols for some well known Lévy models.

**Example 4.6** (Symbol in the Black-Scholes (BS) model). *In the univariate Black-Scholes model, determined by the Brownian volatility  $\sigma >$ , the symbol is given by*

$$A(\xi) = A^{bs}(\xi) = i\xi b + \frac{1}{2}\sigma^2\xi^2, \quad (37)$$

*with drift set  $b$  to*

$$b = r - \frac{1}{2}\sigma^2. \quad (38)$$

**Example 4.7** (Symbol in the Merton model). *In the Merton model where  $\sigma >$ ,  $\lambda > 0$ ,  $\alpha \in \mathbb{R}$  and  $\beta > 0$ , the symbol computes to*

$$A(\xi) = A^{merton}(\xi) = \frac{1}{2}\sigma^2\xi^2 + i\xi b - \lambda \left( e^{-i\alpha\xi - \frac{1}{2}\beta^2\xi^2} - 1 \right) \quad (39)$$

*with drift set to*

$$b = r - \frac{1}{2}\sigma^2 - \lambda \left( e^{\alpha + \frac{\beta^2}{2}} - 1 \right), \quad (40)$$

*as required by the no-arbitrage condition.*

**Example 4.8** (Symbol in the CGMY model). *In the CGMY model of Carr et al. (2002) with  $\sigma > 0$ ,  $C > 0$ ,  $G \geq 0$ ,  $M \geq 0$  and  $Y \in (1, 2)$ , the symbol computes to*

$$A(\xi) = A^{cgmy}(\xi) = i\xi b + \frac{1}{2}\sigma^2\xi^2 - C\Gamma(-Y) \left[ (M + i\xi)^Y - M^Y + (G - i\xi)^Y - G^Y \right], \quad (41)$$

*for all  $\xi \in \mathbb{R}$ , with drift  $b$  set to*

$$b = r - \frac{1}{2}\sigma^2 - C\Gamma(-Y) \left[ (M - 1)^Y - M^Y + (G + 1)^Y - G^Y \right] \quad (42)$$

*for martingale pricing.*

**Example 4.9** (Symbol in the NIG model). *With  $\sigma > 0$ ,  $\alpha > 0$ ,  $\beta \in \mathbb{R}$  and  $\delta > 0$  such that  $\alpha^2 > \beta^2$ , the symbol of the NIG model is given by*

$$A(\xi) = A^{nig}(\xi) = \frac{1}{2}\sigma^2\xi^2 + i\xi b - \delta \left( \sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta - i\xi)^2} \right) \quad (43)$$

for all  $\xi \in \mathbb{R}$  with drift given by

$$b = r - \frac{1}{2}\sigma^2 - \delta \left( \sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + 1)^2} \right) \quad (44)$$

Implementing (36), we encounter new numerical challenges: From the perturbation study in Section 3.3 that we need to evaluate the integrals at high precision. Consider first the Black-Scholes model and choose the piecewise linear hat functions as basis elements as a toy example. Applying a standard Matlab integration routine will lead to considerable errors. To understand the effect, let us have a closer look at the integrands in (36), which we depict in Figure 2.

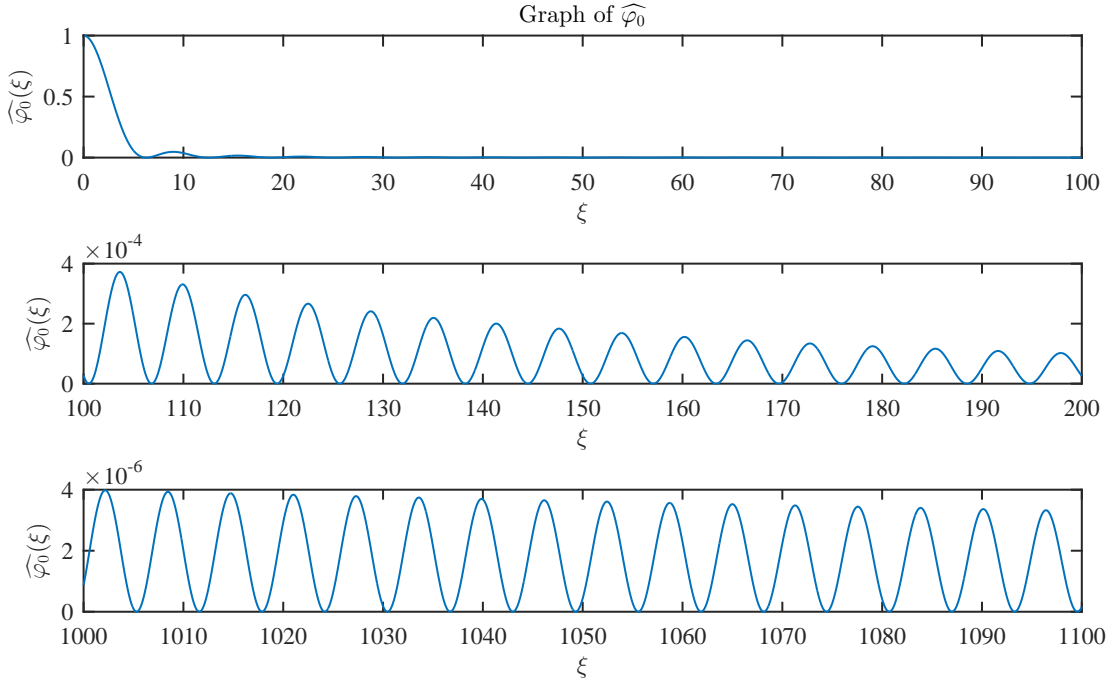


Figure 2: Graph of  $\widehat{\varphi}_0$ , the Fourier transform of the hat function  $\varphi_0$  centered over the origin, evaluated over three subintervals of  $\mathbb{R}^+$ . The mesh is chosen with  $h = 1$ . The oscillations and the rather slow decay to zero complicate numerical integration with high accuracy requirements considerably when  $\widehat{\varphi}_0$  is involved.

More precisely, we show several integrands of  $A \in \mathbb{R}^{N \times N}$  in the representation provided by (36) of Corollary 4.4 with the symbol given by Example 4.6. Each integrand is evaluated for a different value of  $l - k$  over three different subintervals taken from the unbounded integration range. In the Fourier approach of calculating

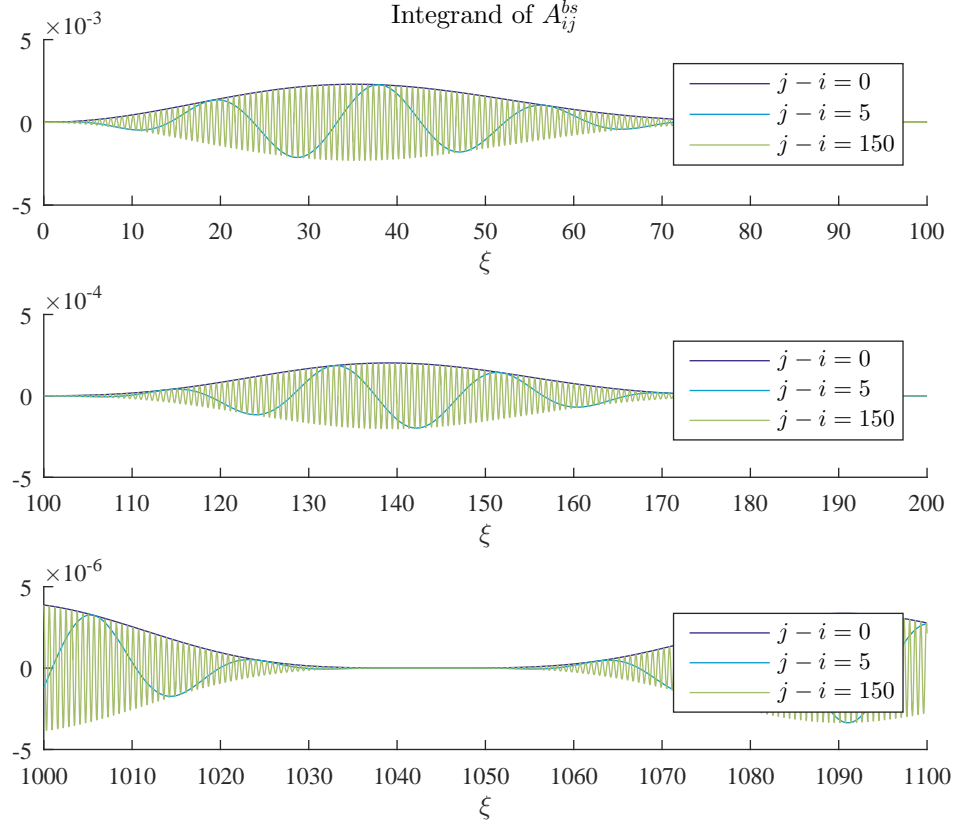


Figure 3: The integrand for the Black-Scholes stiffness matrix  $A_{kl}$  for several values of  $l-k$ . The grid of the hat functions spans the interval  $[-5, 5]$  with 150 equidistantly spaced inner nodes and grid fineness  $h = 0.0662$ . A Black-Scholes solution on this grid would thus be represented by the weighted sum of 150 hat functions. We observe that oscillations of the integrand increase in the value of  $|l-k|$ .



the stiffness matrix  $A \in \mathbb{R}^{N \times N}$  via the respective symbol, the integrands of  $A_{kl}$  would have to be numerically integrated for all  $l - k \in \{-(N-1), \dots, -1, 0, 1, \dots, N-1\}$ .

The larger  $|j-i|$ , however, the more severe the numerical challenges for evaluating the integrand, as Figure 3 demonstrates. All integrands illustrated therein decay rather slowly. Additionally, oscillations increase in  $|j-i|$ . In combination, these two observations seriously threaten a numerically reliable evaluation of the integral. With increasing values of  $|j-i|$ , the oscillations of the integrand accelerate and the number of necessary supporting points for accurate integration soars.

These considerations show us that we need to further investigate the problem to obtain a flexible method to compute the stiffness matrix reliably and with low computational cost. The path that we propose here is to modify the problem in such a way that the resulting integrands decay much faster so that the domain of integration can be chosen considerably smaller and a usual integration routine such as Matlab's function `quadgk` is sufficient. To do so, we first observe that the hat functions, which we used in our toy example, are piecewise linear functions. While being continuous they are not continuously differentiable everywhere and thus lack smoothness on an elementary level already. This lack of smoothness translates into a slow decay of their Fourier transform.

Therefore, we propose to replace the piecewise linear basis functions by basis functions that display considerably higher regularity leading to appealing decay properties of the integrands in (36). In the following two sections, we present two different approaches to implement such a problem modification.

#### 4.1.1 From classic basis functions to mollified hats

It is well known that convolution with a smooth function has a smoothing effect on the function that the convolution is applied to. Our first basis function alternative will therefore be a classic hat function smoothed by convolution. Functions that qualify for this smoothing by convolution are called mollifiers. In order to choose an appropriate mollifier for our purposes—the fast and accurate computation of the integrals in (36), the mollifiers need to display two essential features:

- (1) The Fourier transform of the modified basis function needs to be available.
- (2) The smoothing effect needs to be steerable through a parameter.

As the Fourier transform of the convolution of two functions is the product of the two Fourier transformed functions, (2) boils down to the availability of the Fourier transform of the mollifier. Since the Fourier transform of standard mollifiers is not available in closed form, we widen the range of the standard mollifiers and allow for non-compact support. More precisely, we call the sequence  $m = (m_k)_{k \in \mathbb{N}}$ ,  $m_k \in L^1(\mathbb{R}^d)$  for all  $k \in \mathbb{N}$ , a *mollifier*, if

1.  $m_k \geq 0$ , for all  $k \in \mathbb{N}$ ,
2.  $\int_{\mathbb{R}^d} m_k(x) \, dx = 1$ , and
3. for all  $\varrho > 0$  we have the convergence  $\int_{\mathbb{R}^d \setminus B_\varrho(0)} m_k(x) \, dx \rightarrow 0$  for  $k \rightarrow \infty$ .

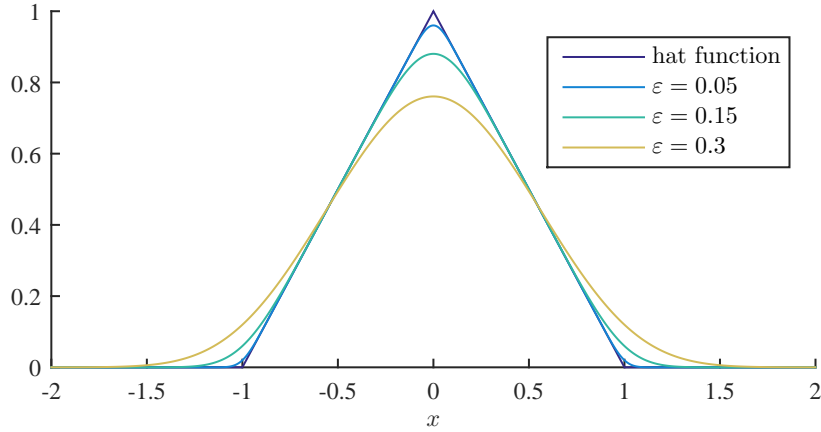


Figure 4: A comparison between the classic hat function  $\varphi_0$  on a grid with  $h = 1$  and the mollified hat function  $\varphi_0^\varepsilon = \varphi_0 * m_{\text{Gaussian}}^\varepsilon$  for several values of  $\varepsilon \in \{0.05, 0.15, 0.3\}$  using the Gaussian mollifier of Example 4.10.

Feature (1) is often required and we follow the usual construction here. By Proposition and Definition 2.14 in Alt (2011) we can adjust the influence of mollification by a mollification  $\varepsilon$ . To this end let  $m \in L^1(\mathbb{R}^d)$  with

$$m \geq 0, \quad \text{and} \quad \int_{\mathbb{R}^d} m(x) dx = 1. \quad (45)$$

Define

$$m^\varepsilon = \frac{1}{\varepsilon^d} m\left(\frac{\cdot}{\varepsilon}\right). \quad (46)$$

Then for each  $\varrho > 0$  we have  $\int_{\mathbb{R}^d} m^\varepsilon(x) dx = 1$  and  $\int_{\mathbb{R}^d \setminus B_\varrho(0)} m^\varepsilon(x) dx \rightarrow 0$  for  $\varepsilon \rightarrow 0$ . Consequently, for each null sequence  $(\varepsilon_k)_{k \in \mathbb{N}}$  the sequence  $(m^{\varepsilon_k})_{k \in \mathbb{N}}$  is a mollifier in the sense of our definition.

**Example 4.10** (A mollifier based on the Normal distribution). *We present an example for a mollifier. Define*

$$m_{\text{Gaussian}}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \quad (47)$$

*Then we call  $(m_{\text{Gaussian}}^{\varepsilon_k})_{k \in \mathbb{N}}$  defined according to (46) a Gaussian mollifier. The characteristic function of the Gaussian mollifier is known in closed form,*

$$\widehat{m_{\text{Gaussian}}^\varepsilon}(\xi) = \exp\left(-\frac{1}{2}\varepsilon^2 \xi^2\right), \quad (48)$$

*thus exhibiting exponential decay.*

It is a well known result, that mollified functions  $f * m_k$  converge to  $f$  in  $L^p(\mathbb{R}^d)$ ,  $1 \leq p < \infty$  when  $k$  tends to infinity, see for example Satz 2.15 in Alt (2011).

Figure 5 displays the integrand in the stiffness matrix of the Black-Scholes model. The integrand is evaluated on three subintervals of the semi-infinite integration

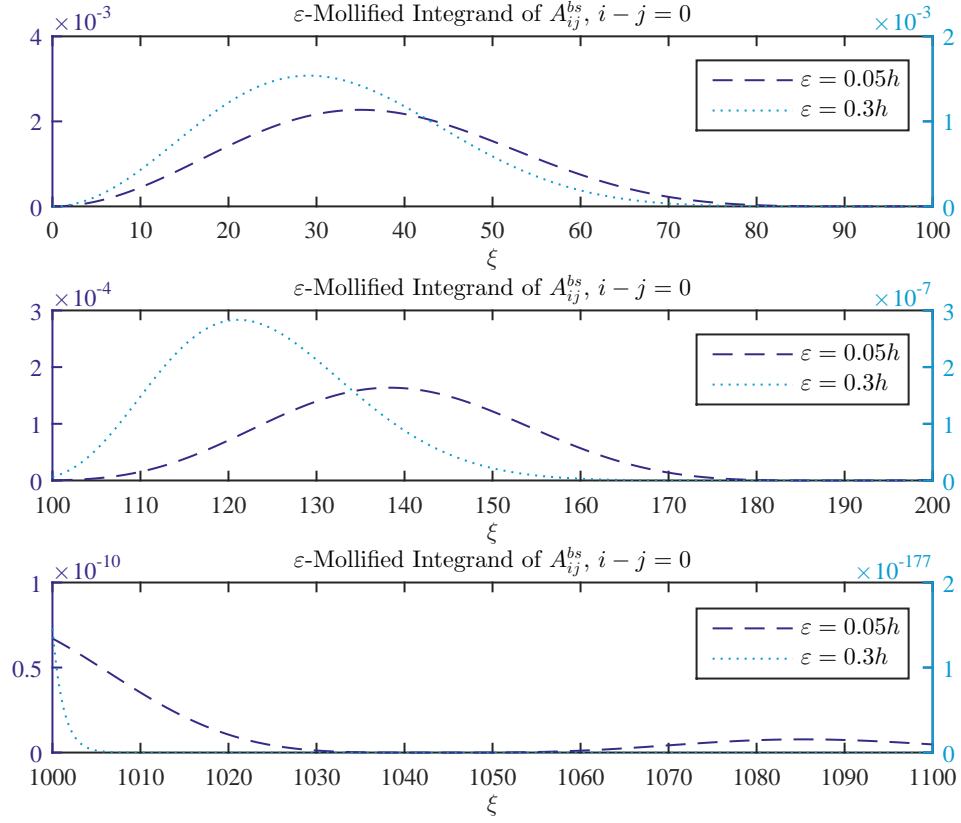


Figure 5: The integrand of  $A_{kl}$ , the stiffness matrix of the Black-Scholes model with mollified hat functions as basis functions for the main diagonal entry,  $l - k = 0$ .

region. The grid setting is identical to the one of Figure 3. Instead of classic hat functions their mollified counterparts have been employed as basis functions using the Gaussian mollifier of Example 4.10 as smoothing influence. Even with just a slight mollification influence,  $\varepsilon = 0.05h$ , the integrand decays noticeably faster thanks to the mollification. For moderate values of  $\varepsilon = 0.3h$  the integrand decays to zero rapidly.

Before we test the performance of this approach to modify the Galerkin scheme in Section 5 below, we introduce our second approach.

#### 4.1.2 Splines as basis functions

Instead of mollification of piecewise linear basis functions, we can choose basis functions that display higher regularity itself. We therefore investigate a well-established class of Finite Element basis functions as candidates for our purposes, namely cubic splines. Spline theory is a well-investigated field that applies to a much broader context than we consider here. We refer the reader to Schumaker (2007) for an introduction and overview. From our perspective, splines are smooth basis functions.

Their Fourier transform is accessible and the theory of function spaces they span is well-established. We give the definition of the Irwin-Hall cubic spline that inherits its name from related probability distribution.

We define the univariate *Irwin-Hall spline*  $\varphi : \mathbb{R} \rightarrow \mathbb{R}^+$  by

$$\varphi(x) = \frac{1}{4} \begin{cases} (x+2)^3 & , -2 \leq x < -1 \\ 3|x|^3 - 6x^2 + 4 & , -1 \leq x < 1 \\ (2-x)^3 & , 1 \leq x \leq 2 \\ 0 & , \text{elsewhere} \end{cases} \quad (49)$$

for all  $x \in \mathbb{R}$ . The spline  $\varphi$  has compact support on  $[-2, 2]$  and is a cubic spline. We use it to define a spline basis:

**Definition 4.1** (Spline basis functions on an equidistant grid). *Choose  $N \in \mathbb{N}$ . Assume an equidistant grid  $\Omega = \{x_1, \dots, x_N\}$ ,  $x_j \in \mathbb{R}$  for all  $j = 1, \dots, N$ , with mesh fineness  $h > 0$ . Let  $\varphi$  the Irwin-Hall spline of (49). For  $j = 1, \dots, N$  define*

$$\varphi_j(x) = \varphi((x - x_j)/h), \quad \forall x \in \mathbb{R}.$$

We call  $\varphi_j$  the spline basis function associated to node  $j$ .

For a given grid  $\Omega = \{x_1, \dots, x_N\}$ ,  $x_j \in \mathbb{R}$ , Definition 4.1 provides the set of spline basis functions that we also use in our numerical implementation, later. In standard spline literature, the set of Irwin-Hall spline basis functions is usually enriched with additional splines associated with the first and the last node of the grid that provide further flexibility in terms of Dirichlet and Neumann boundary conditions, see for example Schumaker (2007). We omit the three Irwin-Hall basis functions associated with either of the first and the last grid nodes thus implicitly prescribing Dirichlet, Neumann and second order derivative zero boundary conditions.

**Lemma 4.11** (Fourier Transform of the Irwin-Hall spline). *Let  $\varphi$  be the Irwin-Hall cubic spline of (49). Then its Fourier transform  $\widehat{\varphi}$  is given by*

$$\widehat{\varphi}(\xi) = \frac{3}{\xi^4} (\cos(2\xi) - 4\cos(\xi) + 3) \quad (50)$$

for all  $\xi \in \mathbb{R}$ .

The proof of the Lemma follows by elementary calculation. This immediately gives the following

**Corollary 4.12** (Fourier Transform of spline basis functions on an equidistant grid). *Choose  $N \in \mathbb{N}$ . Assume an equidistant grid  $\Omega = \{x_1, \dots, x_N\}$ ,  $x_j \in \mathbb{R}$  for all  $j = 1, \dots, N$ , with mesh fineness  $h > 0$  and let  $\varphi_j$  be the spline basis function associated with node  $j$  as defined in Definition 4.1. Its Fourier transform is given by*

$$\widehat{\varphi_j}(\xi) = e^{i\xi x_j} \frac{3}{h^3 \xi^4} (\cos(2\xi h) - 4\cos(\xi h) + 3)$$

for all  $\xi \in \mathbb{R}$ .

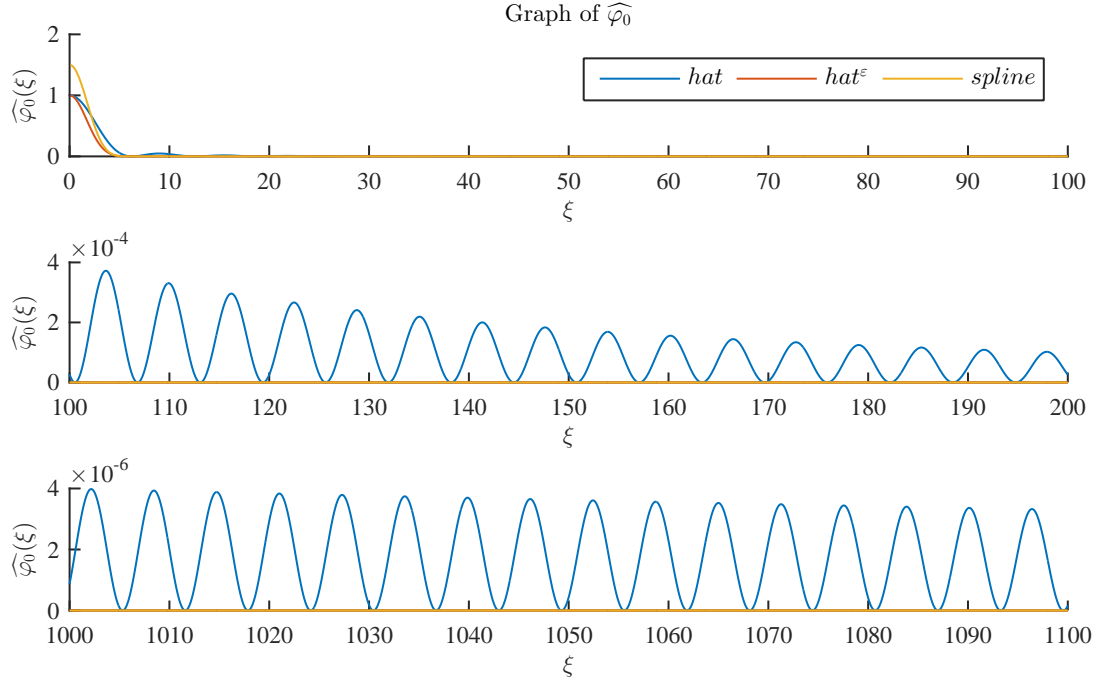


Figure 6: Graphs of the Fourier transforms of all basis function candidates presented in this section, evaluated over three subintervals of  $\mathbb{R}^+$ . The mesh is chosen with  $h = 1$ , the mollification parameter is again set to  $\varepsilon = 0.3h$ .

Figure 6 compares the decay behaviour of the Fourier transforms of all three basis function types. As Figure 2 already illustrated, the Fourier transform of the classic hat functions exhibits both slow decay rates and oscillatory behaviour. In stark contrast the Fourier transforms of the mollified hats as well as the Fourier transform of Irwin-Hall splines visually decays to zero instantly. In case of the mollified hat functions this is due to the exponential decay of the Fourier transform of the Gaussian mollifier while for splines Corollary 4.12 displays a polynomial decay of order 4. In this regard, both alternatives to the classic hat functions are promising candidates for the implementation of the symbol method of Corollary 4.4. In the upcoming Section 5 we put that promise to the test.

## 5 Numerical Implementation

In this section we implement the pricing PIDEs for plain vanilla call and put options and test the two approaches to the symbol method experimentally.

**Theorem 5.1** (Feynman-Kac). *Let  $(L_t)_{t \geq 0}$  be a (time-homogeneous) Lévy process. Consider the PIDE*

$$\begin{aligned} \partial_t V^{C,P} + \mathcal{A}V^{C,P} &= 0, & \text{for almost all } t \in (0, T) \\ V^{C,P}(0) &= g^{C,P}, \end{aligned} \tag{51}$$

where  $\mathcal{A}$  is the operator associated with the symbol of  $(L_t)_{t \geq 0}$ . Assume further the assumptions (A1)–(A3) of Eberlein and Glau (2011) to hold. Then (51) possesses a unique weak solution

$$V^{C,P} \in W^1(0, T; H_\eta^{\alpha/2}(\mathbb{R}^d), L_\eta^2(\mathbb{R}^d)) \quad (52)$$

where  $\alpha > 0$  is the Sobolev index of the symbol of  $(L_t)_{t \geq 0}$  and  $\eta \in \mathbb{R}^d$  is chosen according to Theorem 6.1 in Eberlein and Glau (2011). If additionally  $g_\eta^{C,P} \in L^1(\mathbb{R}^d)$ , then the relation

$$V^{C,P}(T - t, x) = \mathbb{E} [g^{C,P}(L_{T-t} + x)] \quad (53)$$

holds for all  $t \in [0, T]$ ,  $x \in \mathbb{R}^d$ .

*Proof.* The result proved in Eberlein and Glau (2011) and follows from Theorem 6.1 therein.  $\square$

**Remark 5.2.** Setting  $g^{C,P} = g^C$  in (51), the payoff profile of a European call option, results in  $V^C$  being the price of a European call option. Analogously, setting  $g^{C,P} = g^P$ , the payoff profile of a European put option, results in  $V^P$  being the price of a European call option.

Algorithm 1 summarizes the abstract structure of a general FEM solver based on the symbol method. By plugging the symbol associated to the model of choice into the computation of line 9 of the algorithm, the solver instantly adapts to that model. In other words, only one line needs to be specified to obtain a model specific solver for option pricing. As Examples 4.6, 4.7, 4.8, 4.9 and others emphasize, the symbol exists in analytically (semi-)closed form for many models, indeed. Algorithm 1 thus provides a very appealing tool for FEM pricing in practice.

## 5.1 Truncation to zero boundary conditions

As we derive prices of plain vanilla European call and put options, the solution to the respective pricing PIDE is defined on the whole real line. As a first step towards a discretization, we want to truncate the domain to bounded interval  $(a, b)$  and we choose to implement zero boundary conditions. To do so, we follow the standard procedure to subtract an appropriate auxiliary function  $\psi$ . Having chosen  $\psi$ , the resulting modified problem for  $\phi = V^{C,P} - \psi$  is

$$\begin{aligned} \partial_t \phi(t, x) + (\mathcal{A}\phi)(t, x) &= f(t, x), & \forall (t, x) \in (0, T) \times \mathbb{R} \\ \phi(0, x) &= g_\Psi(x), & \forall x \in \mathbb{R}, \end{aligned} \quad (54)$$

where  $g_\Psi(x) = g(x) - \psi(0, x)$  for all  $x \in \mathbb{R}$  and the right hand side  $f$  is given by

$$f(t, x) := -(\partial_t \psi(t, x) + (\mathcal{A}\psi)(t, x)).$$

The solution  $V^{C,P}$  to the original problem (51) can easily be restored by  $V^{C,P} = \phi + \psi$ . We establish the properties that  $\psi$  needs to provide, later, where we will present some examples, as well.

The right hand side in vector notation is given by  $F(t^k) = (F_1(t^k), \dots, F_N(t^k)) \in \mathbb{R}^N$  for each  $t^k$  on the time grid with  $F_j(\cdot)$ ,  $j = 1, \dots, N$ , given by

$$F_j = - \int_{\mathbb{R}} (\partial_t \psi(t, x) + (\mathcal{A}\psi)(t, x) + r\psi(t, x)) \varphi_j(x) dx \quad (55)$$

---

**Algorithm 1** A symbol method based FEM solver

---

- 1: Choose equidistant space grid  $x_i, i = 1, \dots, N$
  - 2: Choose basis functions  $\varphi_i, i = 1, \dots, N$ , with  $\varphi_i(x) = \varphi_0(x - x_i)$  for some  $\varphi_0$
  - 3: Choose equidistant time grid  $T_j, j = 0, \dots, M$
  - 4: **procedure** COMPUTE MASS MATRIX  $M$
  - 5:     Derive the mass matrix  $M \in \mathbb{R}^{N \times N}$  by
  - 6:      $M_{kl} = \int_{\mathbb{R}} \varphi_l(x) \varphi_k(x) dx, \quad \forall k, l = 1, \dots, N$
  - 7: **procedure** COMPUTE STIFFNESS MATRIX  $A$
  - 8:     Derive the stiffness matrix  $A \in \mathbb{R}^{N \times N}$  by plugging the symbol  $A$  of the chosen model into the following formula and computing
  - 9:      $A_{kl} = \frac{1}{2\pi} \int_{\mathbb{R}} A(\xi) e^{i\xi(x_k - x_l)} |\widehat{\varphi_0}(\xi)|^2 d\xi, \quad \forall k, l = 1, \dots, N$
  - 10:    using numerical integration
  - 11: **procedure** RUN THETA SCHEME
  - 12:    Choose a function  $\psi$  to subtract from the original pricing problem to obtain a zero boundary problem and retrieve the respective basis function coefficient vectors  $\overline{\psi}^k \in \mathbb{R}^N, k = 1, \dots, M$ . Consider the suggestions by Lemma 5.3 or Lemma 5.4 for plain vanilla European options below.
  - 13:    Choose an appropriate basis function coefficient vector  $V^1 \in \mathbb{R}^N$  matching the initial condition of the transformed problem
  - 14:    Derive the right hand side vectors  $F^k \in \mathbb{R}^N, k = 0, \dots, M$ . Consult Lemma 5.3 or Lemma 5.4 matching the choice of  $\psi$ .
  - 15:    Choose  $\theta \in [0, 1]$  and run the iterative scheme
  - 16:    **for**  $k = 0 : (M - 1)$
  - 17:        $V^{k+1} = (M + \Delta t \theta A)^{-1} ((M - \Delta t (1 - \theta) A) V^k + F^{k+\theta})$
  - 18:    **end**
  - 19: **procedure** RECONSTRUCT SOLUTION TO ORIGINAL PROBLEM
  - 20:    Add previously subtracted right hand side  $\psi$  to the solution of the transformed problem by computing
  - 21:     $\tilde{V}^k = V^k + \overline{\psi}^k, \quad k = 0, \dots, M$
  - 22:    to retrieve the basis function coefficient vectors  $\tilde{V}^k, k = 0, \dots, M$ , to the original pricing problem
- 

for all  $j = 1, \dots, N$ .

We observe that the operator  $\mathcal{A}$  applied to the auxiliary function  $\psi$  appears in the integral. For the same reasons as in the computation of the stiffness matrix entries, we decide to apply the symbol method for the computation of the entries of the right hand side  $F \in \mathbb{R}^N$ . We pursue these considerations in following section.

## 5.2 Computation of the right hand side $F$

First, we need to choose an appropriate auxiliary function  $\psi$ . As its purpose is to enable us to truncate the domain and insert zero boundary conditions, we need to inspect the limit behaviour of the price value

$$\begin{aligned} V^C(x, t) &\rightarrow 0, & x &\rightarrow -\infty, t \in [0, T] \\ V^C(x, t) &\rightarrow e^x - K e^{-rt}, & x &\rightarrow +\infty, t \in [0, T] \end{aligned} \tag{56}$$

for call options and

$$\begin{aligned} V^P(x, t) &\rightarrow Ke^{-rt} - e^x, & x \rightarrow -\infty, \quad t \in [0, T] \\ V^P(x, t) &\rightarrow 0, & x \rightarrow +\infty, \quad t \in [0, T] \end{aligned} \quad (57)$$

for put options. This is the usual way to obtain the auxiliary function. Now, in regard to our specific approach, relying on the Fourier transforms, we identify additional desirable feature for the auxiliary function. We denote  $\widehat{\psi}(t, z) := \widehat{\psi(t, \cdot)}(z)$ . Consider a smooth function  $\psi : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$  such that  $\psi(t) \in H_{\eta}^{\alpha/2}(\mathbb{R})$  for all  $t \in [0, T]$  for some  $\eta \in \mathbb{R}$ . Then, for the second summand in (55) we have by applying the symbol method of Lemma 4.1 that

$$\int_{\mathbb{R}} (\mathcal{A}\psi)(t, x) \varphi_j(x) dx = \frac{1}{2\pi} \int_{\mathbb{R}} A(\xi - i\eta) \widehat{\psi}(t, \xi - i\eta) \overline{\widehat{\varphi_j}(\xi + \eta)} d\xi, \quad (58)$$

where  $A$  denotes the symbol of the model. With the above identity, we are able to derive the right hand side  $(F_j)_{j=1, \dots, N}$  of (55) in terms of Fourier transforms by

$$F_j = -\frac{1}{2\pi} \int_{\mathbb{R}} \left( \widehat{\partial_t \psi}(t, \xi - i\eta) + A(\xi - i\eta) \widehat{\psi}(t, \xi - i\eta) + r \widehat{\psi}(t, \xi - i\eta) \right) \overline{\widehat{\varphi_j}(\xi + \eta)} d\xi.$$

This shows that  $\psi$  is numerically suitable for the purpose of localizing the pricing PIDE if  $\psi$  is quickly evaluable on the region  $[a, b] \times [0, T]$  and the integrals determining  $F_j$  can be numerically evaluated fast for all  $j = 1, \dots, N$ . The first feature allows retransforming the solution of the localized problem into the solution of the original pricing PIDE, while the second grants the fast numerical derivation of the right hand side in equation (54). These considerations lead us to the following list of desirable features for the auxiliary function  $\psi$  that is required to obey the respective limit conditions (56), (57):

1. a (semi-)closed expression of the function  $\psi$ ,
2. a (semi-)closed expression of its Fourier transform  $\widehat{\psi}$
3. and fast decay of  $|\widehat{\psi}(\xi)|$  and  $|\widehat{\partial_t \psi}(\xi)|$  for  $|\xi| \rightarrow \infty$ .

The smoother  $\psi$ , the faster  $|\widehat{\psi}|$  decays. We thus need different functions  $\psi$  to subtract that not only fulfil the appropriate boundary conditions (56) or (57) but that are also as smooth as possible.

In the following two subsections we will analyze two candidates for  $\psi$  that display the desired features.

A first suggestion for  $\psi$  consists in using Black-Scholes prices as functions in  $x = \log(S_0) \in [a, b]$  and time to maturity  $t \in [0, T]$  for localization of the pricing PIDE. We express the price of a European option with payoff profile  $g^{C,P}$  in the Black-Scholes model in terms of (generalized) Fourier transforms and define  $\psi$  accordingly, as the following Lemma explains.

**Lemma 5.3** (Subtracting Black-Scholes prices). *Choose a Black-Scholes volatility  $\sigma^2 > 0$ . Define  $\psi$  to be the associated Black-Scholes price,*

$$\psi(t, x) = \psi^{bs}(t, x) = e^{-\eta x} e^{-rt} \frac{1}{2\pi} \int_{\mathbb{R}} e^{i\xi x} \widehat{g^{C,P}}(-(\xi + i\eta)) \varphi_{t,\sigma}^{bs}(\xi + i\eta) d\xi, \quad (59)$$



with  $\varphi_{t,\sigma}^{bs}(z) = e^{tA^{bs}(z)}$ . We denote  $A$  the symbol of the associated operator  $\mathcal{A}$ ,  $r \geq 0$  the prevailing risk-free interest rate and choose  $\eta < -1$  and  $\eta > 0$  in the case of a call option and  $\eta$  for the put. Then the right hand side  $F : [0, T] \rightarrow \mathbb{R}^N$  computes to

$$F_j(t) = \frac{1}{2\pi} \int_{\mathbb{R}} \left( (A^{bs} - A)(\xi - i\eta) \right) \widehat{g^{C,P}}(\xi - i\eta) \exp \left( -t \left( r + A^{bs}(\xi - i\eta) \right) \right) \overline{\widehat{\varphi_j}(\xi + i\eta)} d\xi \quad (60)$$

for all  $j = 1, \dots, N$ .

*Proof.* In order to derive the right hand side, we need to represent  $\psi$  in Fourier terms. Since for call and put options,  $\psi \notin L^1(\mathbb{R})$ , we compute the (generalized) Fourier transform of  $\psi$  or the Fourier transform of  $\psi_\eta = e^\eta g$ , respectively. We get

$$\psi_\eta(t, x) = e^{-rt} \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\xi x} \widehat{g^{C,P}}(\xi - i\eta) \varphi_{t,\sigma}^{bs}(-(\xi - i\eta)) d\xi. \quad (61)$$

The integral in (61) is a Fourier (inversion) integral. We read off

$$\widehat{\psi_\eta}(t, \xi) = \widehat{g^{C,P}}(\xi - i\eta) \exp \left( -t \left( r + A^{bs}(\xi - i\eta) \right) \right), \quad (62)$$

where we used the relation between the characteristic function and the symbol of a process. Now,

$$\widehat{\frac{\partial}{\partial t} \psi_\eta}(t, \xi) = - \left( r + A^{bs}(\xi - i\eta) \right) \widehat{\psi_\eta}(t, \xi). \quad (63)$$

Finally, since  $\psi^{bs} \in H_\eta^{\alpha/2}(\mathbb{R})$ , we have that

$$\int_{\mathbb{R}} (\mathcal{A}\psi^{bs})(t, x) \varphi_j(x) dx = \frac{1}{2\pi} \int_{\mathbb{R}} A(\xi - i\eta) \widehat{\psi^{bs}}(t, \cdot)(\xi - i\eta) \overline{\widehat{\varphi_{j-\eta}}(\xi)} d\xi. \quad (64)$$

Collecting our results proves the claim.  $\square$

The candidate  $\psi = \psi^{bs}$  matches the desired criteria. It is quickly evaluable, since functions for yielding Black-Scholes prices are implemented in many code libraries. Also, the integral in (60) is numerically accessible, since the integrand decays fast. Observe that FFT techniques could be employed to compute  $F_j(t)$  for all  $j = 1, \dots, N$  simultaneously. A major disadvantage of choosing  $\psi = \psi^{bs}$ , however, lies in the fact that  $t \in [0, T]$  can not be separated from the integrand in (60). Consequently,  $F_j(t^k)$ ,  $j = 1, \dots, N$ ,  $k = 1, \dots, M$ , must be numerically evaluated on each time grid node individually. This results in significant numerical cost. We therefore present a second candidate for  $\psi$  that avoids this issue.

**Lemma 5.4** (Subtracting Quasi-Hockey stick multiplied by Normal). *Let  $\sigma_\psi > 0$ . Define  $\psi^C$  in the call option and  $\psi^P$  in the put option case by*

$$\begin{aligned} \psi^C(t, x) &= (e^x - Ke^{-rt}) \Phi(x), & (t, x) &\in [0, T] \times [a, b], \\ \psi^P(t, x) &= (Ke^{-rt} - e^x) (1 - \Phi(x)), & (t, x) &\in [0, T] \times [a, b], \end{aligned} \quad (65)$$

where  $\Phi$  denotes the cumulative distribution function of the normal  $\mathcal{N}(0, \sigma_\psi^2)$  distribution. Furthermore, in the call option case choose  $\eta < -1$  and  $\eta > 0$  in the put option case. Then, the right hand side  $F : [0, T] \rightarrow \mathbb{R}^N$  computes to

$$F_j(t) = \frac{1}{2\pi} \left( \int_{\mathbb{R}} (A(\xi - i\eta) + r) \frac{\widehat{f^{\mathcal{N}}}(\xi - i(\eta + 1))}{i\xi + (\eta + 1)} \overline{\widehat{\varphi_j}(\xi + i\eta)} d\xi \right. \\ \left. - e^{-rt} K \int_{\mathbb{R}} (A(\xi - i\eta)) \frac{\widehat{f^{\mathcal{N}}}(\xi - i\eta)}{i\xi + \eta} \overline{\widehat{\varphi_j}(\xi + i\eta)} d\xi \right), \quad (66)$$

for all  $j = 1, \dots, N$  with  $t \in [0, T]$ , where  $A$  is the symbol of the associated operator  $\mathcal{A}$  and with

$$\widehat{f^{\mathcal{N}}}(\xi) = \exp\left(-\frac{1}{2}\sigma_\psi^2 \xi^2\right),$$

the Fourier transform of the normal  $\mathcal{N}(0, \sigma_\psi^2)$  density.

*Proof.* We consider the call option case first. To derive the expression for  $F_j$  in (66) we need to compute the Fourier transform of (the appropriately weighted)  $\psi^C$ . We choose  $\eta < -1$  arbitrary but fix and  $t \in [0, T]$  arbitrary but fix and compute for  $K = 1$ ,

$$\widehat{\psi_\eta^C(t, \cdot)}(\xi) = \int_{\mathbb{R}} e^{i\xi x} e^{\eta x} (e^x - e^{-rt}) \Phi(x) dx \\ = \int_{\mathbb{R}} e^{i\xi x} e^{(\eta+1)x} \Phi(x) dx - e^{-rt} \int_{\mathbb{R}} e^{i\xi x} e^{\eta x} \Phi(x) dx. \quad (67)$$

Integration by parts and l'Hôpital's rule that

$$\int_{\mathbb{R}} e^{i\xi x} e^{(\eta+1)x} \Phi(x) dx = -\frac{1}{i\xi + (\eta + 1)} \int_{\mathbb{R}} e^{i(\xi - i(\eta+1))x} f^{\mathcal{N}}(x) dx, \quad (68)$$

which can be expressed in terms of the Fourier transform of the normal distribution yielding

$$\int_{\mathbb{R}} e^{i\xi x} e^{(\eta+1)x} \Phi(x) dx = -\frac{\widehat{f^{\mathcal{N}}}(\xi - i(\eta + 1))}{i\xi + (\eta + 1)}. \quad (69)$$

Equivalently, we obtain for the second integral in (67) that

$$\int_{\mathbb{R}} e^{i\xi x} e^{\eta x} \Phi(x) dx = -\frac{\widehat{f^{\mathcal{N}}}(\xi - i\eta)}{i\xi + \eta}. \quad (70)$$

Assembling these results we find

$$\widehat{\psi_\eta^C(t, \cdot)}(\xi) = -\frac{\widehat{f^{\mathcal{N}}}(\xi - i(\eta + 1))}{i\xi + (\eta + 1)} + e^{-rt} \frac{\widehat{f^{\mathcal{N}}}(\xi - i\eta)}{i\xi + \eta}. \quad (71)$$

We deduce

$$F_j(t) = \frac{1}{2\pi} \left( \int_{\mathbb{R}} (A(\xi - i\eta) + r) \frac{\widehat{f^{\mathcal{N}}}(\xi - i(\eta + 1))}{i\xi + (\eta + 1)} \overline{\widehat{\varphi_j}(\xi + i\eta)} d\xi \right. \\ \left. - e^{-rt} \int_{\mathbb{R}} A(\xi - i\eta) \frac{\widehat{f^{\mathcal{N}}}(\xi - i\eta)}{i\xi + \eta} \overline{\widehat{\varphi_j}(\xi + i\eta)} d\xi \right) \quad (72)$$

with

$$\widehat{f^{\mathcal{N}}}(\xi) = \exp\left(-\frac{1}{2}\sigma_\psi^2\xi^2\right).$$

For the put option case we choose as defined in (65),

$$\begin{aligned}\psi^P(x, t) &= (Ke^{-rt} - e^x)(1 - \Phi(x)) \\ &= (e^x - Ke^{-rt})(\Phi(x) - 1).\end{aligned}\tag{73}$$

Since

$$\frac{\partial}{\partial x}(\Phi(x) - 1) = \frac{\partial}{\partial x}\Phi(x), \quad \forall x \in \mathbb{R},\tag{74}$$

the computations for  $\widehat{\psi_\eta^P}$  follow along the same lines as they do for the call and we get the relation

$$\widehat{\psi_\eta^P(t, \cdot)}(\xi) = \widehat{\psi_\eta^C(t, \cdot)}(\xi), \quad \forall (t, \xi) \in [0, T] \times \mathbb{R},\tag{75}$$

for  $\eta$  set to some  $\eta > 0$ , which proves the claim.  $\square$

**Remark 5.5** (Computational features of  $\psi^C$  and  $\psi^P$ ). While  $\psi^C$  serves as localizing function for the call option case,  $\psi^P$  can be used in the put option case. Both candidates are based on the payoff functions of call and put options but avoid the lack of differentiability with respect to  $x$  in  $x = \log(Ke^{-rt})$  for  $t \in [0, T]$ . As a consequence, both  $\psi^C$  and  $\psi^P$  are very smooth functions and thus fulfil the requirements collected above when  $\sigma_\psi$  is chosen small enough. Additionally, the two integrals in (66) do not depend on the time variable  $t \in [0, T]$  and thus need to be computed only once for each basis function  $\varphi_j$ . This results in a significant acceleration in computational time compared to the suggestion  $\psi = \psi^{bs}$  of Lemma 5.3.

### 5.3 Empirical Convergence Results

The previous sections have outlined the necessary consecutive phases in setting up a Finite Element solver for option pricing. In order to obtain a flexible FEM solver for pricing PIDEs in Lévy models, the previous section introduced the symbol method which considers all components of the FEM solver in Fourier space. There, components are based on the symbol instead of the Lévy measure and become numerically accessible. Many examples of asset models for which the associated symbols exist in analytically closed form have deemed this alternative approach being worthwhile to pursue. At the same time, however, smoothness of the FEM basis functions became a critical issue. In a second step, we therefore investigated two approaches to combine smoothness and numerical accessibility. Mollified hat functions and splines were introduced as promising examples to construct a symbol method based FEM solver with.

This section will put that promise to the test. We therefore implemented the symbol method for both mollified hats and splines. In stark contrast to a direct implementation of (16) and (55), the symbol method enjoys the flexibility of being able to easily plug in the symbol of any Lévy model for which it is available in analytically closed form. The model restriction of that first approach thus disappears. Having first implemented the method for the Merton model, the extension of the

code to the NIG and the CGMY model comes with virtually no additional effort. In this regard, the method impressively underlines its appeal for applications in practice where the suitability of a model might depend on the asset class it is employed for. An institution that needs to maintain pricing routines for several asset classes will thus cherish the flexibility that the symbol method offers, recall Algorithm 1 in this regard which sketches the implementation of a general, symbol method based FEM solver that easily adapts to various models.

Finally, we conduct an empirical order of convergence study. We consider the univariate Merton, CGMY and NIG model and investigate the empirical rates of convergence for the different implementations as Table 1 summarizes.

Model	Symbol	Parameter choices		Implemented basis functions	
				Mollified hats	Splines
Merton	Example 4.7	$\sigma = 0.15,$ $\beta = 0.2,$	$\alpha = -0.04,$ $\lambda = 3$	✓	✓
CGMY	Example 4.8	$C = 0.5,$ $M = 27.24,$	$G = 23.78,$ $Y = 1.1$	✓	✓
NIG	Example 4.9	$\alpha = 12.26,$ $\delta = 0.52$	$\beta = -5.77,$	✓	✓

Table 1: An overview of the models considered in the empirical order of convergence analysis and their parametrization. For these models, the symbol method is implemented and tested for both mollified hat functions and splines. In addition, we investigate the empirical convergence rate for the Merton model using classic hat functions as basis functions in a classic implementation disregarding the symbol method. In all models, the constant risk-less interest rate has been set to  $r = 0.03$ .

For each model and each implemented basis function type enlisted in Table 1 we conduct an empirical order of convergence study using the pricing problem of a call option with strike  $K = 1$  as an example, thus considering the payoff function

$$g(x) = \max(e^x - 1, 0). \quad (76)$$

In each study we compute FEM prices for  $N_k$  basis functions, with

$$N_k = 1 + 2^k, \quad k = 4, \dots, 9 \quad (77)$$

resulting in  $N_4 = 17$  basis functions in the most coarse and  $N_9 = 513$  basis functions in the most granular case. On each grid, the nodes that basis functions are associated with are equidistantly spaced from another and the basis functions always span the space interval  $\Omega = [-5, 5]$ . The time discretization is kept constant with  $N_{\text{time grid}} = 2000$  equidistantly spaced time nodes spanning a grid range of two years up until maturity, thus covering a time to maturity interval of

$$[T_1, T_{N_{\text{time}}}], \quad \text{with } T_1 = 0 \text{ and } T_{N_{\text{time}}} = 2. \quad (78)$$

For each  $k = 4, \dots, 9$ , the resulting price surface constructed by  $N_k$  basis functions in space and  $N_{\text{time}} = 2000$  grid points in time is computed. A comparison of these surfaces is drawn to a price surface of most granular structure based on the same type of basis function. We call this most granular surface *true* price surface. It rests on  $N_{\text{true}} = N_{11} = 1 + 2^{11} = 2049$  basis functions in space and  $N_{\text{time}}$  grid points in time spanning the same grid intervals as above, that is  $\Omega = [-5, 5]$  in space and  $[0, 2]$  in time, respectively. The underlying FEM implementation is thus based on distances  $h_{\text{true}}$  between grid nodes that basis function are associated with of

$$\begin{aligned} h_{\text{true}}^{\text{mollified hat}} &= (5 - (-5))/(2 + 2^{11}) \approx 0.0049, \\ h_{\text{true}}^{\text{splines}} &= (5 - (-5))/(4 + 2^{11}) \approx 0.0049, \\ \Delta t_{\text{true}} &= 2/(2000 - 1) \approx 0.001 \end{aligned} \tag{79}$$

in space and time, respectively. Note that all space grids are designed in such a way that the log-strike  $\log(K) = 0$  is one of the space nodes. For each model and method and each  $k = 4, \dots, 9$ , the (discrete)  $L^2$  error  $\varepsilon_{L^2}$  is calculated as

$$\varepsilon_{L^2}(k) = \sqrt{\Delta t_{\text{true}} \cdot h_{\text{true}} \cdot \sum_{i=1}^{N_{\text{time}}} \sum_{j=1}^{N_{\text{true}}} \left( \text{Price}_{\text{true}}(i, j) - \text{Price}_k(i, j) \right)^2},$$

wherein  $\text{Price}_{\text{true}}(i, j)$  is the value of the true pricing surface at space node  $j \in \{1, \dots, 1 + 2^{11}\}$  and time node  $i = 1, \dots, 2000$  and  $\text{Price}_k(i, j)$  is the respective, linearly interpolated value of the coarser pricing surface with only  $N_k$  basis function nodes.

Figure 7 summarizes the results of the six studies of empirical order of convergence in the Merton, the NIG and the CGMY model in a symbol based implementation once using mollified hats and once using splines as basis functions. In each implementation and for all considered models, the (discrete)  $L^2$  error decays exponentially with rate 2. The convergence result of Theorem 5.4 by von Petersdorff and Schwab (2003) suggest that this is the best possible rate we can hope for, which yields the experimental validation of both approaches.

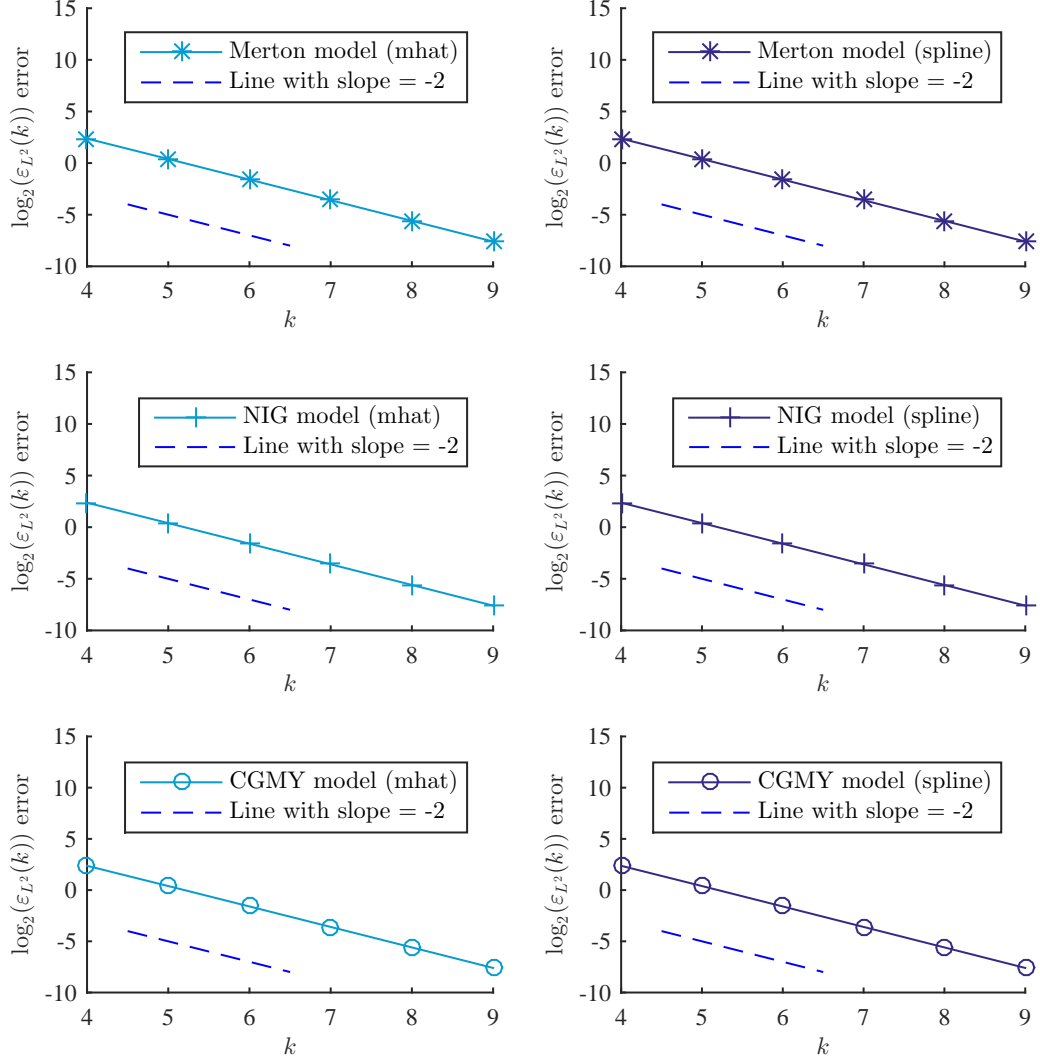


Figure 7: Results of the empirical order of convergence study for the Merton, the NIG and the CGMY model using mollified hats (left pictures) and splines (right pictures) as basis functions. All models are parametrized as stated in Table 1. Additionally, part of a straight line with (absolute) slope of 2 is depicted in each figure serving as a comparison.

## 6 Conclusion and outlook

We have presented a finite element solver that is highly flexible in the model choice and that maintains numerical feasibility. Invoking the symbol was key. The transition into Fourier space has introduced smoothness as a new requirement to the basis functions. We have presented splines and mollified hats as compatible basis functions in our approach. Several numerical examples have confirmed the convergence rates expected by standard theory in both cases. The contribution of the mollification to the error has not yet been addressed in the literature. One possibility to derive such error estimates lies in adapting the perturbation analysis of von Petersdorff and Schwab (2003) treating mollification as a numerical perturbation of the stiffness matrix. Let us mention several possible extensions of the approach. First, the implementation naturally extends to time-inhomogeneous Lévy models that we neglected here for notational convenience. Second, combining the symbol method with Wavelet basis functions allows for compression techniques that might further improve the overall numerical performance, as Hilber et al. (2013) outlines. Third, the polynomial decay that we observe in our numerical experiments can possibly be improved to exponential rates by invoking an *hp*-discontinuous Galerkin scheme, see e.g. Schötzau and Schwab (2006).

## A Proofs

### A.1 Proof of Lemma 3.1

*Proof.* We first consider  $\varphi, \psi \in C_0^\infty(\mathbb{R})$ .

For  $F \equiv 0$  the assertion follows directly from partial integration. Since the Lévy measure may be unbounded around the origin, the representation of the jump part of the bilinear form,

$$a^{jump}(\varphi, \psi) := - \int_{\mathbb{R}} \int_{\mathbb{R}} \left( \varphi(x+y) - \varphi(x) - \varphi'(x) h(y) \right) F(dy) \psi(x) e^{2\langle \eta, x \rangle} dx,$$

needs to be carefully derived. In order to exploit the identity

$$\varphi(x+y) - \varphi(x) - y\varphi'(x) = \int_0^y \int_0^z \varphi''(v) dv dz$$

we split the integral with respect to the Lévy measure in three parts, set  $c(F) := \int_{|y|<1} (y - h(y)) F(dy) - \int_{|y|>1} h(y) F(dy)$  and obtain

$$\begin{aligned} a^{jump}(\varphi, \psi) &:= - \int_{\mathbb{R}} \int_{|y|<1} \int_0^y \int_0^z \varphi''(x+v) dv dz F(dy) \psi(x) e^{2\langle \eta, x \rangle} dx \\ &\quad - c(F) \int_{\mathbb{R}} \varphi'(x) \psi(x) e^{2\langle \eta, x \rangle} dx \\ &\quad - \int_{\mathbb{R}} \int_{|y|>1} (\varphi(x+y) - \varphi(x)) F(dy) \psi(x) e^{2\langle \eta, x \rangle} dx. \end{aligned}$$

Thanks to  $\int_0^y \int_0^z |\varphi''(v)| dv dz \leq cy^2$  with some constant  $c > 0$  for all  $y \in [-1, 1]$  and

$$\begin{aligned} & \int_{\mathbb{R}} \int_{|y|<1} \int_0^y \int_0^z |\varphi'(x+v)| dv dz F(dy) |\psi'(x) + 2\eta\psi(x)| e^{2\langle\eta, x\rangle} dx \\ & \leq (1 + 2\eta) \|\varphi\|_{H_\eta^1} \|\psi\|_{H_\eta^1} \int_{|y|<1} y^2 F(dy) \end{aligned} \quad (80)$$

we can apply the theorem of Fubini and partial integration to obtain

$$\begin{aligned} & - \int_{\mathbb{R}} \int_{|y|<1} \int_0^y \int_0^z \varphi''(x+v) dv dz F(dy) \psi(x) e^{2\langle\eta, x\rangle} dx \\ & = \int_{\mathbb{R}} \int_{|y|<1} \int_0^y \int_0^z \varphi'(x+v) dv F(dy) (\psi'(x) + 2\eta\psi(x)) e^{2\langle\eta, x\rangle} dx. \end{aligned}$$

This yields the assertion for  $\varphi, \psi \in C_0^\infty(\mathbb{R})$ .

Next, we verify that the bilinear form as stated in Lemma 3.1 is well defined for  $\varphi, \psi \in H_\eta^1(\mathbb{R})$  and is continuous with respect to the norm of  $H_\eta^1(\mathbb{R})$ . For  $F \equiv 0$  this is obvious. The assertion follows for the jump part from inequality (80) and

$$\int_{\mathbb{R}} \int_{|y|>1} |\varphi(x+y) - \varphi(x)| F(dy) |\psi(x)| e^{2\langle\eta, x\rangle} dx \leq 2F(\mathbb{R} \setminus [-1, 1]) \|\varphi\|_{L_\eta^2} \|\psi\|_{L_\eta^2}.$$

Thus  $a$  from Lemma 3.1 is a continuous bilinear form on  $H_\eta^1(\mathbb{R}) \times H_\eta^1(\mathbb{R})$  that coincides with (9) on the dense subset  $C_0^\infty(\mathbb{R}) \times C_0^\infty(\mathbb{R})$ . This proves the assertion.  $\square$

## A.2 Proof of Lemma 4.2

*Proof.* To prove the assertion, we verify the conditions of Lemma 3 in Glau (2016b), which provides an abstract robustness result for weak solutions. We first observe that the conditions for  $f_n, f, g_n, g$  coincide with those of Lemma 3 in Glau (2016b). Second, we verify conditions (An1)–(An3) of Lemma 3 in Glau (2016b). Therefore we assign to each  $u, v \in X$  the coefficients  $\alpha_k(u), \alpha_k(v) \in \mathbb{R}$  for  $k \leq N$  such that  $u = \sum_{k=1}^N \alpha_k(u) \varphi_k$  and  $v = \sum_{k=1}^N \alpha_k(v) \varphi_k$ . Thanks to the finite dimensionality of  $X$ , there exists a constant  $\tilde{C} > 0$  such that for all  $u \in X$ ,

$$\|u\|_V \leq \sum_{k=1}^N |\alpha_k(u)| \|u\|_V \leq C' \|u\|_V. \quad (81)$$

Thanks to (27) there exists a sequence  $0 < c_n \rightarrow 0$  such that for all  $j, k \leq N$ ,

$$|(a_n - a)(\varphi_j, \varphi_k)| \leq c_n \|\varphi_j\|_V \|\varphi_k\|_V. \quad (82)$$

Together with assumption (A2) this yields for all  $j, k \leq N$ ,

$$|a_n(\varphi_j, \varphi_k)| \leq C_1 \|\varphi_k\|_V \|\varphi_k\|_V. \quad (83)$$



Inequalities (83) and (81) together yield for all  $u, v \in X$ ,

$$\begin{aligned}
|a_n(u, v)| &\leq \sum_{k=1}^N \sum_{j=1}^N |\alpha_k(u) \alpha_j(v)| |a_n(\varphi_k, \varphi_j)| \\
&\leq C_1 \sum_{k=1}^N \sum_{j=1}^N |\alpha_k(u) \alpha_j(u)| \|\varphi_k\|_V \|\varphi_j\|_V \\
&\leq C_1 \tilde{C}^2 \|u\|_V \|v\|_V,
\end{aligned}$$

which shows that condition (An1) of Lemma 3 in Glau (2016b) is satisfied.

Due to inequalities (82) and (81), we have for all  $u \in X$ ,

$$\begin{aligned}
|(a - a_n)(u, u)| &\leq \sum_{k=1}^N \sum_{j=1}^N |\alpha_k(u) \alpha_j(u)| |a_n(\varphi_k, \varphi_j)| \\
&\leq c_n \sum_{k=1}^N \sum_{j=1}^N |\alpha_k(u) \alpha_j(u)| \|\varphi_j\|_V \|\varphi_k\|_V \\
&\leq c_n \tilde{C}^2 \|u\|_V^2,
\end{aligned}$$

which shows assumption (An3) of Lemma 3 in Glau (2016b).

Finally, from assumption (A1) and the last inequality for all  $u \in X$  we obtain

$$\begin{aligned}
a_n(u, u) &\geq a(u, u) - |(a - a_n)(u, u)| \\
&\geq G \|u\|_V^2 - G' \|u\|_H^2 - c_n \tilde{C}^2 \|u\|_V^2,
\end{aligned}$$

which shows that there exists  $N_0 \in \mathbb{N}$  such that condition (An2) of Lemma 3 in Glau (2016b) is satisfied for all  $n > N_0$ . This shows the assertion of Lemma 4.2.  $\square$

## References

- Alt, H., 2011. Lineare Funktionalanalysis: Eine anwendungsorientierte Einführung, 6th Edition. Springer Lehrbuch. Springer.
- Brummelhuis, R., Chan, R., 2014. An RBF scheme for option pricing in exponential Lévy models. Applied Mathematical Finance 21, 238–269.
- Carr, P., Geman, H., Madan, D. B., Yor, M., 2002. The fine structure of asset returns: An empirical investigation. Journal of Business 75 (2), 305–332.
- Chan, R., Hubbert, S., 2014. Options pricing under the one-dimensional Jump-diffusion model using the radial basis function interpolation scheme. Review of Derivatives Research 17 (2), 161–189.
- Cont, R., Voltchkova, E., 2005a. A finite difference scheme for option pricing in jump diffusion and exponential Lévy models. SIAM Journal on Numerical Analysis 43 (4), 1596–1626.

- Cont, R., Voltchkova, E., 2005b. Integro-differential equations for option prices in exponential Lévy models. *Finance and Stochastics* 9 (3), 299–325.
- Eberlein, E., Glau, K., 2011. PIDEs for pricing European options in Lévy models – a Fourier approach, technische Universität München.
- Eberlein, E., Glau, K., 2014. Variational solutions of the pricing PIDEs for European options in Lévy models. *Applied Mathematical Finance* 21 (5-6), 417–450.
- Föllmer, H., 2009. Alles richtig und trotzdem falsch? *Mitteilungen der DMV* 17, 148–154.
- Glau, K., 2016a. Classification of Lévy processes with parabolic Kolmogorov backward equations, forthcoming in *SIAM Journal Theory of Probability and its Applications*.
- Glau, K., 2016b. Feynman-Kac type formula for Lévy processes with discontinuous killing rate, accepted for publication in *Finance and Stochastics*.
- Hilber, N., Reich, N., Winter, C., Schwab, C., 2009. Numerical methods for Lévy processes. *Finance and Stochastics* 13 (4), 471–500.
- Hilber, N., Reichmann, O., Schwab, C., Winter, C., 2013. *Computational Methods for Quantitative Finance*. Springer.
- Itkin, A., 2015. Efficient solution of backward jump-diffusion PIDEs with splitting and matrix exponentials, forthcoming in *Journal of Computational Finance*.
- Itkin, A., Carr, P., 2012. Using Pseudo-Parabolic and Fractional Equations for Option Pricing in Jump Diffusion Models. *Computational Economics* 40, 63–104.
- Marazzina, D., Reichmann, O., Schwab, C., 2012. hp-DGFEM for Kolmogorov-Fokker-Planck Equations of Multivariate Lévy Processes. *Mathematical Models and Methods in Applied Sciences (M3AS)* 22 (1).
- Matache, A.-M., Nitsche, P.-A., Schwab, C., 2005. Wavelet Galerkin pricing of American options on Lévy driven assets. *Quantitative Finance* 5 (4), 403–424.
- Matache, A.-M., von Petersdorff, T., Schwab, C., 2004. Fast deterministic pricing of options on Lévy driven assets. *Mathematical Modelling and Numerical Analysis* 38 (1), 37–71.
- Salmi, S., Toivanen, J., von Sydow, L., 2014. An IMEX-scheme for pricing options under stochastic volatility models with jumps. *SIAM Journal on Scientific Computing* 36 (4), B817–B834.
- Schötzau, D., Schwab, C., 2006. Time Discretization of Parabolic Problems by the *hp*-Version of the Discontinuous Galerkin Finite Element Method. *SIAM Journal of Numerical Analysis* 38 (3), 837–875.
- Schumaker, L., 2007. *Spline Functions: Basic Theory*, 3rd Edition. Cambridge Mathematical Library. Cambridge University Press.

- von Petersdorff, T., Schwab, C., 2003. Wavelet Discretizations of Parabolic Integro-differential Equations. *SIAM Journal on Numerical Analysis* 41 (1), 159–180.
- Wloka, J., 1987. *Partial Differential Equations*. Cambridge University Press.
- Zeidler, E., 1990. *Nonlinear Functional Analysis and its Applications*. Vol. II/A. Springer, Linear Monotone Operators.